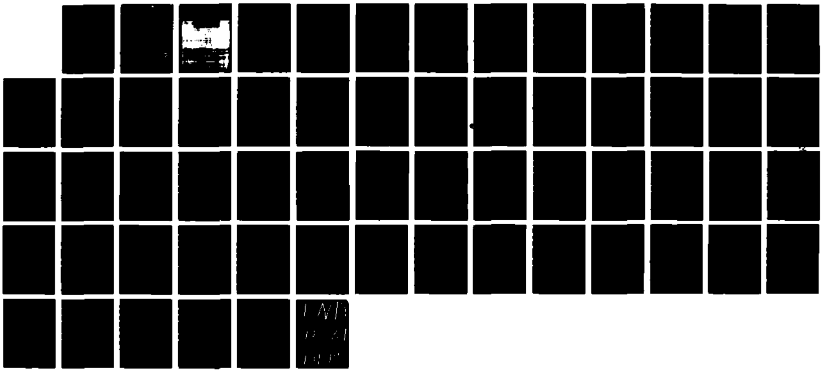
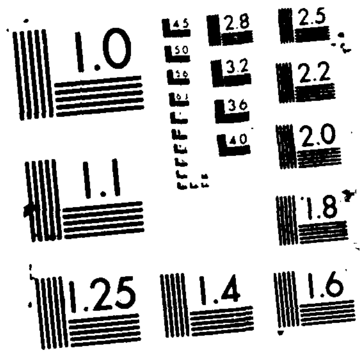
AD-A185 805 OPTIMAL AND APPROXIMATELY OPTIMAL CONTROL POLICIES FOR QUEUES IN HEAVY TR..(U) BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMICAL SYSTE.. 1/1

UNCLASSIFIED H J KUSHNER ET AL.. MAR 87 LCDS/CCS-87-24 F/G 12/4 NL




1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

DTIC FILE COPY

①

SECURITY CLASS

TION PAGE

Form Approved
OMB No. 0704-0188

AD-A185 805

1a. REPORT SEC	1b. RESTRICTIVE MARKINGS
2a. SECURITY C	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 87-1517
6a. NAME OF PERFORMING ORGANIZATION / 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR-TR- 87-1517
6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION / 8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-85-0315
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

12. PERSONAL AUTHOR(S)

13a. TYPE OF REPORT	13b. TIME COVERED FROM ______ TO ______	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

DTIC
ELECTE
NOV 0 6 1987
S D
E

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code) / 22c. OFFICE SYMBOL

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

AFOSR-TR- 87-1517

OPTIMAL AND APPROXIMATELY OPTIMAL CONTROL POLICIES FOR QUEUES IN HEAVY TRAFFIC

by

H.J. Kushner and K.M. Ramachandran

March 1987 LCDS/CCS #87-24

IN DEO SPERAMUS

Lefschetz Center for Dynamical Systems
and
Center for Control Sciences

Division of Applied Mathematics Brown University Providence RI 02912

87 14 419

OPTIMAL AND APPROXIMATELY OPTIMAL CONTROL POLICIES FOR QUEUES IN HEAVY TRAFFIC

by

H.J. Kushner and K.M. Ramachandran

March 1987 LCDS/CCS #87-24

OPTIMAL AND APPROXIMATELY OPTIMAL CONTROL POLICIES

FOR QUEUES IN HEAVY TRAFFIC

by

Harold J. Kushner*

and

K.M. Ramachandran**

Lefschetz Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, Rhode Island 02912

March 1987

* The work of this author was supported in part by the National Science Foundation under contract #ECS-8505674, the Air Force Office of Scientific Research under contract #AFOSR-85-0315, and the Office of Naval Research under contract #N00014-85-K-0607.

** The work of this author was supported in part by the Army Research Office under contract #DAAG29-84-K-0082 and the Office of Naval Research under contract #N00014-85-K-0607.

Abstract

We treat the 'approximately' optimal control problem for tandem queueing or production networks (with local feedback allowed) under heavy traffic. The buffers (scaled with traffic) are finite. The controls allow various inputs, connecting links and the processors to be shut down or opened, in order to manage the system. The service and arrival rates, as well as the routing probabilities can also be controlled, and the system statistics can depend on the system state (scaled buffer occupancies). The associated costs involve holding costs, costs for shutting off/on the links or processors and the opportunity cost for lost production. It is shown that the (scaled) controlled system converges weakly (in an appropriate sense) to a controlled limit 'reflected' diffusion. In the rescaled time, the actions of the controllers lead to multiple 'simultaneous' impulses in the limit problem. Thus we have a non-standard limit control problem, and the usual methods of weak convergence for systems under heavy traffic must be modified. Since the optimal or nearly optimal controls for the physical process are usually not possible to get, it is of considerable interest to know whether an optimal or nearly optimal control for the limit process is also nearly optimal for the physical system with heavy traffic. This is shown to be true, under reasonable conditions. Although the limit control problem is non-standard and there is little available theory concerning it, acceptable numerical procedures are available.

Key Words: Weak convergence, queueing networks, production networks, heavy traffic approximations, controlled reflected diffusions, controlled queueing networks, approximately optimal stochastic controls, numerical methods for stochastic control.

I. Introduction

We consider optimal and 'nearly optimal' control problems for the open queueing networks in heavy traffic of the type dealt with in the fundamental papers of Reiman [1] and Harrison, [2], [3]. Owing to the state and control dependence (in our problem) of the routing, arrival and service time processes, as well as to our use of finite buffers, and to some approximations which are used in [1] - [3] in the modelling in these papers, much of their methodology cannot be carried over. We do try to retain their structure and results wherever possible. One of the main motivations behind the heavy traffic approximations [1] - [4] of queueing networks is the idea that the limit process (which is a reflected Brownian motion in the past work, and a more general impulsively or singularly controlled reflected diffusion here) is easier to analyze than the actual physical process, and that *it is much easier to find good or optimal control policies for the limit than for the physical process*. This is undoubtably true, particularly *if* the traffic is truly heavy the buffer size large or if the routing parameters and input and service times are correlated or state (queue size) dependent.

In [1], one has several interconnected service or processing stations, and at each there is an infinite buffer (ours is finite, but suitably scaled with traffic intensity). At each there are possible arrivals from outside the network as well as arrivals routed from other service stations. The departures are routed (perhaps randomly) to other service stations (perhaps to one that they had previously visited) or leave the network. Eventually (w.p.1) all customers leave the network. Under reasonable conditions on the interarrival and service times

and with appropriate spacial and temporal normalizations, in the heavy traffic case the vector of the normalized queue lengths (the normalized number in the buffers plus in service) converges weakly to a reflected Brownian motion with constant drift and covariance parameters [1]. This will be generalized here in several directions, although we work with a somewhat simpler network structure.

Although it underlies a lot of the motivation for the limit theorems, there seems to have been very little work on the usefulness of the limit process for purposes of getting a good or nearly optimum control for the physical process. Let ϵ index the traffic intensity. As $\epsilon \to 0$, the 'intensity' goes to ∞. For whatever cost criteria is used (this will be defined in later sections), let $V^{\epsilon}(\pi)$ denote its value for the physical system when a policy π is used. Suppose that $\bar{\pi}^{\epsilon}$ is an 'adaptation' of the optimal (or δ-optimal) policy for the limit, applied *to the physical process. (We will say more* about such adaptations later.) For $\bar{\pi}^{\epsilon}$ to be a 'good' policy for the physical process we need at least that $V^{\epsilon}(\bar{\pi}^{\epsilon}) - \inf_{\pi} V^{\epsilon}(\pi)$ be small for small ϵ, where the inf is over an appropriate set of policies for the physical process. This is the problem addressed here. In the course of the development, a number of interesting and non-classical problems arise; for example, the appropriate 'limit' control problem might involve multiple 'simultaneous' impulses, and we must treat state dependent service, arrival and routing processes.

There are many possibilities for the structure of the control problem. Ours, to be described below, illustrates the main problems and develops a (weak convergence based) method which applies to many other formulations. We are forced to differ in several important respects from models used in earlier work

on the limit theorems for queueing networks in heavy traffic. If the service or arrival rates can be controlled, then the limit process is no longer a reflected Brownian motion with constant coefficients; we wish to allow these rates to depend on the system state; we must deal with (implicitly or explicitly) a dynamically controlled upper bound to the buffer size. (Even if the buffer size is infinite, the optimal control might force it to be shut down); owing to the control, there might be 'travel' along the boundaries; some controls (e.g., on/off controls with associated impulsive costs) might yield nice process paths in 'real' time but in the usual interpolated time (i.e., for the sequence for which we seek the weak convergence) the paths between the on/off times move faster and faster as $\epsilon \to 0$ and converge to a discontinuity - but not in the Skorohod topology; the nature of the convergence at these discontinuities can yield (an interesting) limit process with 'multiple simultaneous impulses'; the lumping together of all idle times as done in [1, eqn(3)] in the $B_k(t)$ argument is a useful 'approximation', but it is inappropriate in our context owing to the state and control dependencies, and is not quite the exact physical model in any case (although it yields the correct results); to show that the 'limit' controls and other quantities are 'admissible', or non-anticipative with respect to the limit Brownian motions or reflected diffusions, we need an approach that is at least partly along the lines of the martingale method. In fact, we combine the ideas of [1] with those of the martingale method and the weak convergence techniques of [5], [6].

The work here is a continuation of the lines of development in [6], [7], [8] where approximations to other optimal control problems are dealt with. Owing to the special features of the controlled heavy traffic network of queues, this

past work is not applicable to this problem without major change. We refer to it where helpful in simplifying or reducing an argument.

In Section 2, the basic system is described, the control problem defined and assumptions stated. Many of the results are true for controlled networks allowing general feedback as in [1]. But, in order to avoid some quite complicated bookkeeping, we eventually specialize to a tandem case - with only two processors and feedback only allowed from a processor to itself. The general results can be readily extended to problems where (except for the possibility of rerouting an output back to the input of the same processor), the flow is all 'forward'. In Section 3, we discuss representations for the processes which facilitate the weak convergence analysis, and in Section 4, we describe the proper 'limit' control problem (and some of its peculiarities), i.e., the appropriate controlled reflected diffusion whose optimal (or δ-optimal) controls are to be used for the physical process.

Section 5 contains the basic weak convergence results, and we state and prove the results concerning the 'almost optimality' of the δ-optimal (for small δ) controls for the limit process, when applied to the physical process. Some computational questions are discussed in Section 6. Although the 'limit' control problem is not always simple, effective and convenient numerical methods are available.

2. Problem Description and Assumptions

We start by describing a network with K service stations (processors), the i^{th} referred to as P_i. Each processor services only one customer at a time (although, as will be seen from the development in the sequel, batch or multi server cases can all be handled and even controlled. Shortly, we specialize to the case $K = 2$, but it is simpler to first use a unified terminology. We retain the basic interconnection structure of [1], but use a discrete time parameter for notational simplicity. Each processor can be connected to an external input as well as receive (and deliver) outputs from (to) other processors.

Let $\{\alpha_n^{i,\epsilon}\}$ denote the sequence of interarrival times of the customers coming from the exterior of the network directly to P_i, and let $\xi_n^{i,\epsilon}$ denote the indicator of the event that there was an arrival from the exterior to P_i at time n. As is frequently done (e.g., as *in [1]*), we *adapt the convenient* representation where the processor keeps processing even if the queue is empty, with the 'errors' generated by this convention accounted for by an added *reflection term. With this convention in mind,* let $\{\Delta_n^{i,\epsilon}\}$ denote the sequence of service times for P_i, and $\psi_n^{i,\epsilon}$ the indicator of the event that a service at P_i is completed at time n (whether or not there are actual 'physical' customers in P_i at that time). As in [1], we suppose that if there is an arrival to P_i in the midst of a service interval when the queue at P_i is empty, then the actual service time for that customer is just the residual service time for the current service interval. Under the heavy traffic assumption, this does not affect the limit formulas. Let $I_n^{ij,\epsilon}$, $i = 1,\cdots, K$, $j = 0,\cdots, K$, denote the indicator function of the event that a completed service at P_i at time n

is scheduled to be sent to P_j (or to the exterior, if $j = 0$). We use $\{p_{ij}, i,j = 1,\cdots,K\}$ to denote the probability that a completed service from P_i is to be routed to P_j, and write $p_{i0} = 1 - \sum_{j=1}^{K} p_{ij}$. The buffer size at P_i is $B_i/\sqrt{\epsilon}$, for $B_i > 0$.

The allowable control efforts are as follows. We work with impulsive controls only, although the results can be extended to the case where the service and interarrival 'rates' as well as the routing probabilities are controlled contiuously. The processor P_i can be shut off for a time, at a cost $k_i > 0$, to be paid *at the moment of shut off*. The external inputs to P_i can be shut off for a time, at a cost $k_{0i} > 0$, *to be paid at the moment of shut off*. If P_i communicates to P_j, in lieu of shutting P_i off, we can open or break the link connecting P_i to P_j. In that case the output of P_i which is destined for P_j will be shunted to the exterior and lost, or sold as a 'partially completed' product. The cost for shutting the link off is $k_{ij} > 0$, *to be paid at the moment of shut off*, and there will be an additional cost for the lost customers. This cost is $q_{ij}\sqrt{\epsilon}$ per lost customer, $q_{ij} > 0$. By convention, we allow all customers in P_i who have completed service there and are destined to return to P_i immediately to do so. If the buffer of P_i is full, then one or more inputs must be turned off, i.e., either the input links to P_i are shunted to the exterior, or the P_j connecting to P_i are shut off.

The bulk of the work will use the above control possibilities. The extension to the case where the marginal service or external arrival rates (or even the routing probabilities) are controlled is not a difficult extension and is discussed at the end of the paper.

Let $P_n^{i,\epsilon}$, $P_n^{0i,\epsilon}$ and $P_n^{ji,\epsilon}$, resp., denote the indicators of the events that

P_i is working at time n (i.e., processing or not shut off), the external input to P_i is not shut off at time n, and the link connecting P_j to P_i is open at time n, resp. Let $N_n^{i,\epsilon}$ (resp., $\tilde{N}_n^{i,\epsilon}$) denote the n^{th} time that P_i is turned off (resp., turned back on), and set $\tilde{N}_0^{i,\epsilon} = 0$. Let $N_n^{ij,\epsilon}$ ($i = 0,1,\cdots,K$, $j = 1,\cdots,K$) (resp., $\tilde{N}_n^{ij,\epsilon}$) denote the n^{th} time that the link connecting P_i to P_j is shut off (turned back on, resp.) (If $i = 0$, then it's for the link connecting the exterior to P_j.) Define $v_n^{i,\epsilon} = \epsilon N_n^{i,\epsilon}$, $v_n^{ij,\epsilon} = \epsilon N_n^{ij,\epsilon}$, and similarly define $\tilde{v}_n^{i,\epsilon}$ and $\tilde{v}_n^{ij,\epsilon}$.

Let $X_n^{i,\epsilon} = \sqrt{\epsilon}$ [Number of customers in or waiting for service at P_i at time n] and set $X^{i,\epsilon}(t) = X_{t/\epsilon}^{i,\epsilon}$. This is the quantity of interest in the desired interpolated time and amplitude scale. Then, in this interpolated scale, $[v_n^{i,\epsilon},\tilde{v}_n^{i,\epsilon})$, $n \geq 1$, etc., are the intervals of closure of P_i, etc. When ratios t/ϵ are used as indices, we use the integral part. Until Sections 5 and 6, w.l.o.g., and for notational convenience, we always assume that all processors and links are working at $t = 0$. Thus $\tilde{v}_0^{\alpha} \equiv 0$ and $\tilde{v}_n^{\alpha,\epsilon} > v_n^{\alpha,\epsilon}$ for $n > 0$. In general, it is possible that $v_1^{\alpha,\epsilon} = 0$ also (instantaneous change in the system at the starting time). The optimal value function will depend on the initial system configuration, and the true state of the system is actually the pair (X_n^{ϵ}, status of links and processors). We return to this in Section 5.

In order to keep track of the flows in the system for purposes of the control problem and the limit theorems, we need to separate out the corrections to the flows due to empty queues and to the flow components due to the control actions. *Throughout the paper, ϵ-superscripts will be omitted in the terms in sums or integrals.* The subscript c is for 'combined', since we use it when there is a condition on the status of two controls simultaneously. Define

(2.1a) $$Y^{ij,\epsilon}(t) = \sqrt{\epsilon} \sum_{n=1}^{t/\epsilon} \psi_n^i \, I_n^{ij} \, P_n^i \, I_{\{X_n^i=0\}}$$

(2.1b) $$U^{ij,\epsilon}(t) = \sqrt{\epsilon} \sum_{n=1}^{t/\epsilon} \psi_n^i \, I_n^{ij} \, (1 - P_n^i), \; i \neq 0, \; j \neq i.$$

(2.2a) $$U^{0i,\epsilon}(t) = \sqrt{\epsilon} \sum_{n=1}^{t/\epsilon} \xi_n^i \, (1 - P_n^{0i}), \; i \neq 0,$$

(2.2b) $$U_c^{ji,\epsilon}(t) = \sqrt{\epsilon} \sum_{n=1}^{t/\epsilon} \psi_n^j \, I_n^{ji} (1 - P_n^{ji} \, P_n^j), \; j \neq i, \; j \neq 0,$$

(2.2c) $$Y_c^{ji,\epsilon}(t) = \sqrt{\epsilon} \sum_{n=1}^{t/\epsilon} \psi_n^j \, I_n^{ji} \, P_n^j \, P_n^{ji} \, I_{\{X_n^j=0\}}, \; j \neq i, \; j \neq 0.$$

(2.3) $$A^{i,\epsilon}(t) = \sqrt{\epsilon} \sum_{n=1}^{t/\epsilon} \xi_n^i, \quad D^{ij,\epsilon}(t) = \sqrt{\epsilon} \sum_{n=1}^{t/\epsilon} \psi_n^i \, I_n^{ij}, \; i \neq 0.$$

(2.4) $$Z^{ij,\epsilon}(t) = \sqrt{\epsilon} \sum_{n=1}^{t/\epsilon} \psi_n^i \, I_n^{ij} (1 - P_n^{ij}) \, P_n^i \, I_{\{X_n^i \neq 0\}}, \; i \neq j, \; i,j \neq 0.$$

With the definitions (2.1) to (2.3), we can write

(2.5) $$\begin{aligned} X^{i,\epsilon}(t) = {} & A^{i,\epsilon}(t) + \sum_{j\neq i} D^{ji,\epsilon}(t) - \sum_{j\neq i} D^{ij,\epsilon}(t) \\ & + \sum_{j\neq i} Y^{ij,\epsilon}(t) - \sum_{j\neq i} Y_c^{ji,\epsilon}(t) - U^{0i,\epsilon}(t) \\ & + \sum_{j\neq i} U^{ij,\epsilon}(t) - \sum_{j\neq i} U_c^{ji,\epsilon}(t). \end{aligned}$$

The first term in (2.5) represents the potential external arrivals to P_i, the second represents potential arrivals from other P_j, $j \neq i$, all neglecting the effects of controls or empty queues. The third term represents potential departures from P_i, again neglecting the effects of controls or empty queues. The other terms correct for these omissions. The $Y^{ij,\epsilon}(\cdot)$ corrects for departures from P_i when P_i is working and its queue is empty, and the

$Y_c^{ji,\epsilon}(\cdot)$ corrects for arrivals to P_i from P_j when the buffer of P_j is empty and neither P_j nor the link from P_j to P_i is shut off. The $U^{0i,\epsilon}(\cdot)$ corrects for the stopped external arrivals, when the input to P_i from the exterior is shut off. The $U^{ij,\epsilon}(\cdot)$ corrects for the stopped departures from P_i when P_i is closed, and the $U_c^{ji,\epsilon}(\cdot)$ corrects for the stopped arrivals from P_j to P_i when either P_j is not working or the link from P_j to P_i is shut off (i.e., shunted to the exterior).

The $Z^{ij,\epsilon}(\cdot)$ represents the lost output when the link from P_i to P_j is shunted to the exterior. There can only be lost output at time n if $X_n^{i,\epsilon} > 0$ and $P_n^{i,\epsilon} = 1$ (as well as $P_n^{ij,\epsilon} = 0$). Write $X^\epsilon = (X^{1,\epsilon},\dots,X^{K,\epsilon})$ and let π^ϵ or π denote control policies (i.e., rules for determining the $v^{i,\epsilon}$, $\tilde{v}^{i,\epsilon}$, $v^{ij,\epsilon}$, $\tilde{v}^{ij,\epsilon}$), and let E_x^π denote the expectation, given policy π and initial condition $X_0^\epsilon = x$. Let P denote the *vector* of *indicator functions* $\{P^\alpha\}$ of the processors and links. *In general, the value function depends on* the initial value of P (although we set (w.l.o.g.) the initial values $P_0^\alpha = 1$ until Section 5). Then, for a bounded and continuous $k(\cdot)$ and $\beta > 0$, our cost will be of the discounted form (2.6).

$$
\begin{aligned}
(2.6)\qquad V^\epsilon(\pi,x,P) &= E_x^\pi \int_0^\infty e^{-\beta t}\, k(X^\epsilon(t))dt \\
&+ E_x^\pi \sum_{i=1}^K k_i \sum_n e^{-\beta v_n^{i,\epsilon}} \\
&+ E_x^\pi \sum_{i=0}^K \sum_{j=1}^K k_{ij} \sum_n e^{-\beta v_n^{ij,\epsilon}} \\
&+ E_x^\pi \int_0^\infty e^{-\beta t} \Big[\sum_i q_{0i}\, dU^{0i,\epsilon}(t) + \sum_{i,j=1}^K q_{ij}\, dZ^{ij,\epsilon}(t) \Big]
\end{aligned}
$$

The first term in (2.6) is the holding cost. The next two are the costs for the impulsive switching, and the last the cost of lost output via either non-admittance of customers or forcing them out of the system before the total required processing is completed.

The average cost per unit time problem could be handled as well, but is somewhat more complicated. See, for example the average cost per unit time problems in [6], [8], for other models.

We now specialize to the case of Figure 2.1. We specialize since it is awkward to keep track of the effects of the controls in a network with general feedback allowed, particularly of the effects of empty queues which are (at least partly) due to the control actions. With mainly notational changes, the case dealt here with can be extended to the general case where the only allowed feedback in the system is from the output of a processor to its own input - otherwise the flow is 'forward'.

Refer to Figure 2.1, and assume (A2.1). The first part of this assumption (or restriction on the control actions) says simply that if a queue is empty, then we won't continue to 'starve' it - but will turn on all the inputs. The assumption seems to be quite unrestrictive, and it does simplify the bookkeeping quite a bit

A2.1. *If* $X_n^{2,\epsilon} = 0$, *then all inputs to* P_2 *are open; i.e.,* $P_n^{1,\epsilon} = P_n^{12,\epsilon} = P_n^{02,\epsilon} = 1$. *If* $X_n^{1,\epsilon} = 0$, *then the input to* P_1 *is open* (i.e., $P_n^{01,\epsilon} = 1$). *If some* $X_n^{i,\epsilon} = B_i$, *then all inputs to* P_i *are closed.*

For the system of Figure 2.1, and under (A2.1), we have that (2.1) -

(2.5) take the forms (2.7) - (2.9). Here, $P_n^{2,\epsilon} \equiv 1$, since there is never a need to shut P_2 off. ((2.7) is written for easy reference; all the Y^{ij}, U^{ij} are still defined by (2.1) - (2.2).)

$$Y_c^{12,\epsilon}(t) = \sqrt{\epsilon} \sum_{1}^{t/\epsilon} \psi_n^1 I_n^{12} P_n^1 P_n^{12} I_{\{X_n^1=0\}} \tag{2.7}$$

$$Y^{20,\epsilon}(t) = \sqrt{\epsilon} \sum_{1}^{t/\epsilon} \psi_n^2 I_n^{20} I_{\{X_n^2=0\}}$$

$$Z^{12,\epsilon}(t) = \sqrt{\epsilon} \sum_{1}^{t/\epsilon} \psi_n^1 I_n^{12} (1 - P_n^{12}) P_n^1 I_{\{X_n^1\neq 0\}}$$

$$= U_c^{12,\epsilon}(t) - U^{12,\epsilon}(t) - \sum_{1}^{\infty} \int_{v_n^{12}\cap t}^{\bar{v}_n^{12}\cap t} dY^{12,\epsilon}(s)$$

The $Y^{12,\epsilon}(\cdot)$ will converge to a continuous function and $\bar{v}_n^{12,\epsilon} - v_n^{12,\epsilon} \xrightarrow{\mathcal{L}} 0$. Thus the last term on the right of the last equation will disappear in the limit. Define $U^{1,\epsilon}(\cdot) = U^{10,\epsilon}(\cdot) + U^{12,\epsilon}(\cdot)$. Then

$$X^{1,\epsilon}(t) = A^{1,\epsilon}(t) - D^{10,\epsilon}(t) - D^{12,\epsilon}(t) + Y^{10,\epsilon}(t) + Y^{12,\epsilon}(t) - U^{01,\epsilon}(t) + U^{1,\epsilon}(t) \tag{2.8a}$$

$$X^{2,\epsilon}(t) = A^{2,\epsilon}(t) - D^{20,\epsilon}(t) + D^{12,\epsilon}(t) + Y^{20,\epsilon}(t) - Y_c^{12,\epsilon}(t) - U^{02,\epsilon}(t) - U_c^{12,\epsilon}(t) \tag{2.8b}$$

$$\begin{aligned} V^{\epsilon}(\pi,x,P) &= E_x^{\pi} \int_0^{\infty} e^{-\beta t} k(X^{\epsilon}(t))dt \\ &+ k_1 E_x^{\pi} \sum_n e^{-\beta v_n^{1,\epsilon}} + \sum_{i=1}^{2} k_{0i} E_x^{\pi} \sum_n e^{-\beta v_n^{0i,\epsilon}} \\ &+ k_{12} E_x^{\pi} \sum_n e^{-\beta v_n^{12,\epsilon}} \\ &+ E_x^{\pi} \int_0^{\infty} e^{-\beta t} [\sum_{i=1}^{2} q_{oi} dU^{0i,\epsilon}(t) + q_{12}\, dZ^{12,\epsilon}(t)]. \end{aligned} \tag{2.9}$$

We now give some more definitions and state the heavy traffic assumptions. It will sometimes be convenient to write the multiple sequence $v^\epsilon \equiv \{v_n^{i,\epsilon},\ \bar{v}_n^{i,\epsilon},\ v_n^{ij,\epsilon},\ \bar{v}_n^{ij,\epsilon}\}$ as a single sequence. Let $\{\tau_n^\epsilon\}$ denote the sequence of event *times* indicated by all the elements of v^ϵ in order of increasing time, but without respect to which events they indicate, or whether they indicate multiple events. Define $R_n^\epsilon = (R_n^{1,\epsilon},\ R_n^{01,\epsilon},\ R_n^{02,\epsilon},\ R_n^{12,\epsilon})$, where $R_n^{\alpha,\epsilon} = 1, -1$ or 0 depending on whether or not the 'control' with the same superscript was opened (turned on), closed (turned off) or left unchanged at τ_n^ϵ. *From* $\{R_n^\epsilon, \tau_n^\epsilon\}$, *we can recover all the control actions and their times.*

Let $S_{a,n}^{i,\epsilon} = \sum_{j=1}^{n} \alpha_j^{i,\epsilon}$, $S_{d,n}^{i,\epsilon} = \sum_{1}^{n} \Delta_j^{i,\epsilon}$. Let $E_{a,n}^{i,\epsilon}$ denote the expectation given the arrival, departure and control intervals and actions which ended by real time $S_{a,n}^{i,\epsilon}$, as well as the lengths of all other arrival and service intervals (other than $\alpha_{n+1}^{i,\epsilon}$) which started by but which might not have been completed by time $S_{a,n}^{i,\epsilon}$. Analogously, that $E_{d,n}^{i,\epsilon}$ denote the expectation given the arrival, departure and control intervals and actions which ended by real time $S_{d,n}^{i,\epsilon}$, as well as the lengths of all other arrival and service intervals (other than $\Delta_{n+1}^{i,\epsilon}$) which started by $S_{d,n}^{i,\epsilon}$. Define the conditional variances $\mathrm{var}_{a,n}^{i,\epsilon}$, $\mathrm{var}_{d,n}^{i,\epsilon}$ analogously. Define

$$E_{a,n}^{i,\epsilon}\, \alpha_{n+1}^{i,\epsilon} = \bar{\alpha}_{n+1}^{i,\epsilon},\ \mathrm{var}_{a,n}^{i,\epsilon}\, \alpha_{n+1}^{i,\epsilon} = (\sigma_{a,n+1}^{i,\epsilon})^2$$
$$E_{d,n}^{i,\epsilon}\, \Delta_{n+1}^{i,\epsilon} = \bar{\Delta}_{n+1}^{i,\epsilon},\ \mathrm{var}_{d,n}^{i,\epsilon} \Delta_{n+1}^{i,\epsilon} = (\sigma_{d,n+1}^{i,\epsilon})^2.$$

Henceforth when we say that P_i, P_{0i} *or* P_{12}, *resp., is open (closed) at time* n, *we mean that processor* i *is working, the link from the exterior to* P_i *is open or (resp.), the link from* P_1 *to* P_2 *is open for traffic.*

We will use

A2.2. *There are positive numbers* g_{ai} *and* g_{di} *and bounded continuous functions* $a^i(\cdot)$ *and* $d^i(\cdot)$ *such that*

$$[\overline{\alpha}^{i,\epsilon}_{n+1}]^{-1} = g_{ai} + \sqrt{\epsilon}\, a_{in} + o(\sqrt{\epsilon}),$$

$$[\overline{\Delta}^{i,\epsilon}_{n+1}]^{-1} = g_{di} + \sqrt{\epsilon}\, d_{in} + o(\sqrt{\epsilon}),$$

where $a_{in} = a^i(X^\epsilon_{S^{i,\epsilon}_{a,n}})$ *and* $d_{in} = d^i(X^\epsilon_{S^{i,\epsilon}_{d,n}})$

Comment on (A2.2). We allow the (marginal) external inter-arrival intervals and the service intervals to depend on the system state. The argument $X^\epsilon_{S^{i,\epsilon}_{a,n}}$ (for example) is the proper one, since $S^{i,\epsilon}_{a,n}$ is the (real) starting time for the $(n+1)^{st}$ (external) inter-arrival interval to P_i (the moment of arrival to P_i of the $n+1^{st}$ customer from the outside), and $X^\epsilon_{S^{i,\epsilon}_{a,n}}$ is the system state at that time. We could let the marginal mean rates $a^i(\cdot)$ and $d^i(\cdot)$ be controlled. We then use $a^i(X^i_{S^{i,\epsilon}_{a,n}}, r^{i,\epsilon}_{S^{i,\epsilon}_{a,n}})$, etc. Here the r^α_β is the control over the mean marginal rate. There is no problem in incorporating controlled rates into the weak convergence and approximation results of Section 5. An appropriate associated cost would include a direct cost (higher for higher rates) and an indirect cost due to the possible gain in production due to the higher (input) rates. Similarly, the $g_{\alpha i}$ can be controlled or even state dependent, provided only that the heavy traffic assumption (A2.4) below continues to hold.

A2.3. *The set* $\{|\alpha^{i,\epsilon}_n|^2, |\Delta^{i,\epsilon}_n|^2, i, n < \infty$, *small* ϵ, *all control actions} is uniformly integrable.*

A2.4. *(Heavy traffic assumption)*

$$g_{a1} = (1 - p_{11})g_{d1}$$

$$[p_{12}g_{d1} + g_{a2}]/(1-p_{22}) = g_{d2}.$$

(A2.4) is also what one would get from Reiman's [1] formulas for the case of Figure 2.1. If either condition in (A2.4) is violated, then either some buffer will always be full as $\epsilon \to 0$ (and the cost will go to ∞) or else some $X^{i,\epsilon}(t) \to 0$ as $\epsilon \to 0$ (and the cost will go to ∞). With little extra trouble it is possible to control the p_{ij} also - but this seems to be of not much interest for the case of Figure 2.1. The results for our case can readily be extended to the case of 'feedforward' systems, where the only allowed feedback in the routing is from a processor to itself. For these general cases, it might be worth controlling (marginally) the p_{ij}. The extension is simple, and follows the same lines as would the extension to marginally controlled rates.

A2.5. *The routing variables* $\{I_k^{ij,\epsilon}, i,j,k\}$ *are mutually independent and independent of the* $\{\alpha_k^{i,\epsilon}, \Delta_k^{i,\epsilon}\}$ *and* $P\{I_k^{ij,\epsilon} = 1\} = p_{ij}$.

A2.6. There are continuous functions $\sigma_{ai}(\cdot)$, $\sigma_{di}(\cdot)$ *such that*

$$\sigma_{a,n+1}^{i,\epsilon} = \sigma_{a,i}(X^{\epsilon}_{g^{i,\epsilon}_{a,n}}) + \delta'_{\epsilon}$$

$$\sigma_{d,n+1}^{i,\epsilon} = \sigma_{d,i}(X^{\epsilon}_{g^{i,\epsilon}_{d,n}}) + \delta''_{\epsilon}.$$

where $\delta^{\alpha}_{\epsilon} \xrightarrow{\epsilon} 0$, *uniformly in all other variables.*

Comment on (A2.5) and (A2.6). We allow the conditional variance to

depend on the state here, just to show the possibilities. Controlled variances can also be handled. In many applications (and in most past works on the heavy traffic model) the $\sigma_{\alpha,i}$ are just constants. The independence in (A2.5) can also be weakened, and the sequence of interarrival times or service intervals can be correlated (in ways other than via the 'state' dependence used here). This would involve a more complex method for obtaining the weak convergence. The perturbed test function methods of [5] (see also [6]) are quite suitable for that task, and would require only moderate changes in the proof of Theorem 5.1, but the additional notational, etc, burden seems hardly worth it now.

3. A Convenient Representation for $X^{\epsilon}(\cdot)$.

In this section, we center and rewrite the terms of (2.8), so as to facilitate the weak convergence analysis in Section 5. We will do three things. First, the A and D processes will be centered, the centering terms simplified, and the centered processes written as a rescaling of simpler processes. This is similar to the procedure of [1]. Then we will represent the $Y^{ij,\epsilon}$ and $U^{ij,\epsilon}$ in terms of simpler processes $\bar{Y}^{i,\epsilon}$ and $\bar{U}^{i,\epsilon}$ (not depending on j) plus a term which will go to zero as $\epsilon \to 0$. Finally, we will represent $\bar{Y}^{i,\epsilon}$ and $X^{i,\epsilon}$ as continuous (and unique) functions of the 'other' data, similar to the representation used in [1].

Centering of the Arrival and Departure Processes. Now, several processes will be defined. Define $\bar{S}_a^{i,\epsilon}(t)$ (and analogously $\bar{S}_d^{i,\epsilon}(t)$) to be the *inverse* of the interpolated arrival time function $\epsilon S^{i,\epsilon}_{d,t/\epsilon}$ in the sense that

$$\bar{S}_a^{i,\epsilon}(t) = \max\{\epsilon k:\ \epsilon\, S^{i,\epsilon}_{a,k} \le t\}.$$

Define the centered processes

$$\tilde{A}_0^{i,\epsilon}(t) = \sqrt{\epsilon}\sum_{k=1}^{t/\epsilon}\ \sum_{\ell=S^{i,\epsilon}_{a,k}}^{S^{i,\epsilon}_{a,k+1}-1}\left[\xi^i_\ell - \frac{1}{\bar{\alpha}^i_k}\right]$$

$$= \sqrt{\epsilon}\sum_{1}^{t/\epsilon}(1 - \alpha^i_k/\bar{\alpha}^i_k),$$

(3.1)

$$\tilde{D}_0^{ij,\epsilon}(t) = \sqrt{\epsilon}\sum_{k=1}^{t/\epsilon}\ \sum_{\ell=S^{i,\epsilon}_{d,k}}^{S^{i,\epsilon}_{d,k+1}-1}\left[\psi^i_\ell\, I^{ij}_\ell - \frac{p_{ij}}{\bar{\Delta}^i_k}\right]$$

$$= \sqrt{\epsilon}\sum_{k=1}^{t/\epsilon}\left[I^{ij}_{S^i_{d,k}} - p_{ij}\Delta^i_k/\bar{\Delta}^i_k\right].$$

The second equality in the first definition follows from the fact that $\xi_\ell^{i,\epsilon} = 1$ only at the left endpoint in the interval $[S_{a,k}^{i,\epsilon}, S_{a,k+1}^{i,\epsilon})$ and the length of the interval is $\alpha_k^{i,\epsilon}$ (and similarly for the second definition).

Owing to the independence assumptions in (A2.5), we can (and will, henceforth) replace the $I^{ij}_{S^i_{d,k}}$ by I_k^{ij}. We can write $A^{i,\epsilon}(\cdot)$ in the form (which defines $\tilde{A}^{i,\epsilon}(\cdot)$ and $\tilde{B}_a^{i,\epsilon}(\cdot)$)

$$(3.2) \qquad A^{i,\epsilon}(t) = \sqrt{\epsilon} \sum_{\epsilon k=1}^{\bar{S}_a^{i,\epsilon}(t)} \sum_{\ell = S_{a,k}^{i,\epsilon}}^{S_{a,k+1}^{i,\epsilon}-1} [\xi_\ell^i - \frac{1}{\bar{\alpha}_k^i}] + \sqrt{\epsilon} \sum_{\epsilon k=1}^{\bar{S}_a^{i,\epsilon}(t)} \alpha_k^i/\bar{\alpha}_k^i$$

$$\equiv \tilde{A}_0^{i,\epsilon}(\bar{S}_a^{i,\epsilon}(t)) + \tilde{B}_a^{i,\epsilon}(t) \equiv \tilde{A}^{i,\epsilon}(t) + \tilde{B}_a^{i,\epsilon}(t).$$

Doing the same thing for the $D^{ij,\epsilon}(\cdot)$, we have (which defines $\tilde{D}^{ij,\epsilon}(\cdot)$ and $\tilde{B}_d^{ij,\epsilon}(\cdot)$)

$$(3.3) \qquad D^{ij,\epsilon}(t) = \tilde{D}_0^{ij,\epsilon}(\bar{S}_d^{i,\epsilon}(t)) + \tilde{B}_d^{ij,\epsilon}(t) \equiv \tilde{D}^{ij,\epsilon}(t) + \tilde{B}_d^{ij,\epsilon}(t)$$

where

$$(3.4) \qquad \tilde{B}_d^{ij,\epsilon}(t) = \sqrt{\epsilon} \sum_{\epsilon k=1}^{\bar{S}_d^{i,\epsilon}(t)} \frac{\Delta_k^i}{\bar{\Delta}_k^i} p_{ij}.$$

For purposes of calculation below, write

$$\tilde{D}_0^{10,\epsilon}(t) + \tilde{D}_0^{12,\epsilon}(t) = \sqrt{\epsilon} \sum_{1}^{t/\epsilon} [(1 - I_k^{11}) - (1 - p_{11}) \frac{\Delta_k^i}{\bar{\Delta}_k^i}]$$

We now cancel the 'principal parts' of the $\tilde{B}_\alpha^{i,\epsilon}$ terms. By taking the terms in the order in which they would appear in the centering of the first three terms of(2.8a) and using the expansion in (A2.2), we write

$$\tilde{B}_a^{1,\epsilon}(t) - (\tilde{B}_d^{10,\epsilon}(t) + \tilde{B}_d^{12,\epsilon}(t)) =$$

(3.5)
$$\sqrt{\epsilon} \sum_{\epsilon k=1}^{\bar{S}_a^{1,\epsilon}(t)} \alpha_k^1 [g_{a1} + \sqrt{\epsilon}\, a_{ik} + o(\sqrt{\epsilon})]$$

$$- \sqrt{\epsilon} \sum_{\epsilon k=1}^{\bar{S}_d^{1,\epsilon}(t)} \Delta_k^1 [g_{a1} + \sqrt{\epsilon}\, d_{1k} + o(\sqrt{\epsilon})](1 - p_{11}).$$

Since $\sum_{\epsilon k=1}^{\bar{S}_a^{1,\epsilon}(t)} \alpha_k^1 = t/\epsilon$ (mod. $0(1)$), the principal term of the first sum is $g_{a1}t/_\epsilon$ (mod $0(\sqrt{\epsilon})$), and of the second in $g_{d1}t/_{\sqrt{\epsilon}}$ (mod $0(\sqrt{\epsilon})$). These cancel by (A2.4). By using the definitions of α_k^1 and a_{1k} and the fact that X_k^ϵ changes by at most $O(\epsilon)$ per step, we can write the sum of the middle terms in the first sum of (3.5) as

$$\epsilon \sum_1^{t/\epsilon} a^1(X_k) \quad + \quad (\text{term which} \rightarrow 0 \text{ as } \epsilon \rightarrow 0)$$

and similarly for the analogous terms in the second sum.

With the above cancellations and the last representation, we can rewrite (3.5) as (3.6) (where $\delta_\epsilon^i(\cdot) \xrightarrow{\mathcal{E}} 0$, uniformly on bounded intervals). Equation (3.6) defines $B^{1,\epsilon}(\cdot)$ and $b^1(\cdot)$.

(3.6)
$$\epsilon \sum_1^{t/\epsilon} [a^1(X_k) - (1 - p_{11})d^1(X_k)] + \delta_\epsilon^1(t)$$
$$\equiv \int_0^t b^1(X^\epsilon(s))ds + \delta_\epsilon^1(t) \equiv B^{1,\epsilon}(t) + \delta_\epsilon^1(t).$$

Repeating the procedure for the 'biases' arising from (2.8b), we get (which defines $B^{2,\epsilon}(\cdot)$ and $b^2(\cdot)$)

$$\text{(3.7)} \qquad \tilde{B}_a^{2,\epsilon}(t) - \tilde{B}_d^{20,\epsilon}(t) + \tilde{B}_d^{12,\epsilon}(t)$$

$$= B_2^{2,\epsilon}(t) + \delta_\epsilon^2(t) = \int_0^t b^2(X^\epsilon(s))ds + \delta_\epsilon^2(t)$$

$$= \epsilon \sum_1^{t/\epsilon} [a^2(X_k) - (1 - p_{22})d^2(X_k) + p_{12}d^1(X_k)] + \delta_\epsilon^2(t)$$

$$= \int_0^t b^2(X^\epsilon(s))ds + \delta_\epsilon^2(t) = B^{2,\epsilon}(t) + \delta_\epsilon^2(t).$$

A Representation for $U^{ij,\epsilon}, Y^{ij,\epsilon}$. Define the processes (with $P_n^{2,\epsilon} \equiv 1$)

$$\text{(3.8)} \qquad Y^{i,\epsilon}(t) = \sqrt{\epsilon} \sum_1^{t/\epsilon} \psi_n^j P_n^i I_{\{X_n^j = 0\}}$$

We can also write

$$\text{(3.9a)} \qquad U^{12,\epsilon}(t) = \sum_{n=1}^\infty \int_{v_n^{1,\epsilon} \cap t}^{\bar{v}_n^{1,\epsilon} \cap t} dU_c^{12,\epsilon}(s),$$

$$\bar{Y}_c^{12,\epsilon}(t) = \sum_{n=0}^\infty p_{12} \int_{\bar{v}_n^{12,\epsilon}}^{v_{n+1}^{12,\epsilon}} d\bar{Y}^{1,\epsilon}(s).$$

It will turn out (Section 5) that the limits in (3.9b) hold

$$Y^{1j,\epsilon}(\cdot) - p_{12}Y^{1,\epsilon}(\cdot) \Rightarrow 0$$

$$Y^{20,\epsilon}(\cdot) - (1 - p_{22})Y^{2,\epsilon}(\cdot) \Rightarrow 0$$

$$Y_c^{12,\epsilon}(t) - \sum_{n=0}^\infty p_{12} \int_{\bar{v}_n^{12,\epsilon} \cap t}^{v_{n+1}^{12,\epsilon} \cap t} dY^{1,\epsilon}(s) \Rightarrow 0$$

$$Y_c^{12,\epsilon}(\cdot) - p_{12}Y^{1,\epsilon}(\cdot) \Rightarrow 0$$

$$\text{(3.9b)} \qquad U^{1j,\epsilon}(\cdot) - U^{1,\epsilon}(\cdot)p_{1j}/(1 - p_{11}) \Rightarrow 0, \ j = 0,2$$

$$U_c^{12,\epsilon}(t) - \sqrt{\epsilon}\, p_{12} \sum_0^{t/\epsilon} \psi_n^1(1 - P_n^1 P_n^{12}) \Rightarrow 0$$

$$Z^{12,\epsilon}(\cdot) - [U_c^{12,\epsilon}(\cdot) - U^{12,\epsilon}(\cdot)] \Rightarrow 0$$

$$\bar{v}_n^{\alpha,\epsilon} - v_n^{\alpha,\epsilon} \to 0, \text{ each } \alpha, n.$$

In order to prepare for the utilization of these convergences and simplifications. rewrite (2.8) and (2.9) as follows, where the $\rho^{i,\epsilon}(\cdot)$ and $\hat{\rho}^{i,\epsilon}$ are 'small error' processes and the $W^{i,\epsilon}(\cdot)$ are defined to be the sum of the first three terms in the middle part of (3.10a) and (3.10b), resp.

$$\begin{aligned} X^{1,\epsilon}(t) &= \tilde{A}^{1,\epsilon}(t) - (\tilde{D}^{10,\epsilon}(t) + \tilde{D}^{12,\epsilon}(t)) + B^{1,\epsilon}(t) \\ &\quad + (Y^{10,\epsilon}(t) + Y^{12,\epsilon}(t)) - U^{01,\epsilon}(t) + U^{1,\epsilon}(t) + \rho^{1,\epsilon}(t) \\ &= W^{1,\epsilon}(t) + B^{1,\epsilon}(t) + (1-p_{11})Y^{1,\epsilon}(t) \\ &\quad - U^{01,\epsilon}(t) + U^{1,\epsilon}(t) + \hat{\rho}^{1,\epsilon}(t) \end{aligned} \tag{3.10a}$$

$$\begin{aligned} X^{2,\epsilon}(t) &= \tilde{A}^{2,\epsilon}(t) - \tilde{D}^{20,\epsilon}(t) + \tilde{D}^{12,\epsilon}(t) + B^{2,\epsilon}(t) \\ &\quad + Y^{20,\epsilon}(t) - Y_c^{12,\epsilon}(t) - U^{02,\epsilon}(t) - U_c^{12,\epsilon}(t) + \rho^{2,\epsilon}(t) \\ &= W^{2,\epsilon}(t) + B^{2,\epsilon}(t) + (1-p_{22})Y^{2,\epsilon}(t) - p_{12}Y^{1,\epsilon}(t) \\ &\quad - U^{02,\epsilon}(t) - U_c^{12,\epsilon}(t) + \hat{\rho}^{2,\epsilon}(t) \end{aligned} \tag{3.10b}$$

(3.11) $V^{\epsilon}(\pi,x) =$ [eqn (2.9) with $Z^{12,\epsilon}(\cdot)$ replaced by $U_c^{12,\epsilon}(\cdot) - U^{12,\epsilon}(\cdot)$ and an 'error' term $\hat{\rho}^{3,\epsilon}(\cdot)$ added].

It will turn out (Section 5) that, for any sequence of controls π^{ϵ} with $\sup V^{\epsilon}(\pi^{\epsilon},x) < \infty$, $\sup_{t\le T} |\rho^{i,\epsilon}(t)| \to 0$ in distribution for any $T < \infty$, and similarly for the $\hat{\rho}^{i,\epsilon}(\cdot)$.

Owing to the impulsive nature of the 'control' part of the cost (2.9), on any bounded time interval there are only a finite number (w.p.1) of subintervals on which the controls are active (i.e., where some P_i or P_{ij} is shut off). By the definitions, the reflection terms $Y^{ij,\epsilon}(\cdot)$ cannot increase on these 'control intervals'. In particular, $Y^{1,\epsilon}(\cdot)$ (and $Y^{1i,\epsilon}(\cdot)$) can only increase wh[illegible] [illegible]h P_{01} and P_1 are on (recall that P_{01} is on when $X^1 = 0$). Also, $Y^{2,\epsilon}(\cdot)$ (and

$Y^{20,\epsilon}(\cdot))$ can increase only when all of P_1, P_{12} and P_{02} are on (by (A2.1), if $X_n^{2,\epsilon} = 0$, then all inputs must be turned on). Because of this, the setup of [1, Lemma 1] can be used to obtain the 'reflection' terms as continuous functions of the other 'non-control' data, simply by using the representation of [1] on the appropriate 'non-control' time segments, and we now formalize this.

Let $J_n^{1,\epsilon} = [\mu_n^{1,\epsilon}, \bar{\mu}_n^{1,\epsilon})$ denote the sequence of successive intervals (of interpolated time) such that $P_k^{1,\epsilon} = P_k^{01,\epsilon} = 1$ for $\epsilon k \in J_n^{1,\epsilon}$, and let $J_n^{2,\epsilon} = [\mu_n^{2,\epsilon}, \bar{\mu}_n^{2,\epsilon})$ denote the successive intervals such that $P_k^{1,\epsilon} = P_k^{12,\epsilon} = P_k^{02,\epsilon} = 1$ for $\epsilon k \in J_n^{2,\epsilon}$. The $Y^{i,\epsilon}(\cdot)$ can increase only on the $J_n^{i,\epsilon}$.

We can use the representation for the $Y^{ij,\epsilon}$ of [1] in the pieces between the control intervals. For any function $f(\cdot)$ define $f_{i,n}(\cdot) = f((\mu_n^{i,\epsilon} + \cdot) \cap \bar{\mu}_n^{i,\epsilon})) - f(\mu_n^{i,\epsilon})$. By [1, Lemma 1], there is a unique continuous function $\tilde{F}(\cdot) = (\tilde{F}^1(\cdot), \tilde{F}^2(\cdot))$ (the continuity in the arguments which are functions is taken to be continuity in the topology of uniform convergence on bounded time intervals) such that

$$
\begin{aligned}
(3.12) \qquad & Y_{1,n}^{10,\epsilon} + Y_{1,n}^{12,\epsilon} = \tilde{F}^1(X^{1,\epsilon}(\mu_n^{1,\epsilon}),\ W_{1,n}^{1,\epsilon}(\cdot),\ B_{1,n}^{1,\epsilon}(\cdot),\ \rho_{1,n}^{1,\epsilon}(\cdot)) \\
& Y_{2,n}^{20,\epsilon} = \tilde{F}^2(X^{2,\epsilon}(\mu_n^{2,\epsilon}),\ W_{2,n}^{2,\epsilon}(\cdot),\ B_{2,n}^{2,\epsilon}(\cdot),\ Y_{2,n}^{10,\epsilon}(\cdot) + Y_{2,n}^{12,\epsilon}(\cdot),\ \rho_{2,n}^{2,\epsilon}(\cdot)).
\end{aligned}
$$

Furthermore $\tilde{F}(\cdot)$ is 'non-anticipative', the corresponding $X^{i,\epsilon}(\cdot)$ is non-negative and the (resp.) left hand sides of (3.12) can increase only at those times when the (resp.) $X_n^{i,\epsilon}(\cdot)$ are zero.

Alternatively, there is a unique continuous function $F(\cdot)$ such that

$$
\begin{aligned}
(3.13) \qquad & (Y^{10,\epsilon}(\cdot) + Y^{12,\epsilon}(\cdot),\ Y^{20,\epsilon}(\cdot)) = \\
& F(W^{\epsilon}(\cdot),\ B^{\epsilon}(\cdot),\ \rho^{\epsilon}(\cdot),\ X_0^{\epsilon},\ X^{i,\epsilon}(\mu_n^{i,\epsilon}),\ \mu_n^{i,\epsilon}, \bar{\mu}_n^{i,\epsilon},\ i = 1,2,\ n < \infty)
\end{aligned}
$$

where $F(\cdot)$ has the properties ascribed to $\tilde{F}(\cdot)$ above. In particular, the value of the left side of (3.13) at time t depends only on the arguments of the functions in F at times $\leq t$ and on the $\mu_n^{i,\epsilon}$, $\tilde{\mu}_n^{i,\epsilon}$ with values less than or equal to t. Owing to (3.13), we will not need to concern ourselves with the weak convergence of the arguments of the $Y^{ij,\epsilon}(\cdot)$. This will follow from the weak convergence of the arguments of $F(\cdot)$.

A Tentative Form for the Limit Control Problem. Purely formally, let the arguments of $F(\cdot)$ converge to $W(\cdot)$, $B(\cdot)$, μ_n^i, $\tilde{\mu}_n^i$, $(\rho(\cdot) = 0)$ and let $Y^i(\cdot)$ be the limit of $Y^{i,\epsilon}(\cdot)$. Then, on each bounded time interval the complement of $\{[\mu_n^i,\tilde{\mu}_n^i),n < \infty\}$ will just be a finite set of points, and the controls will be impulses acting at these points. Using this assumed convergence and (3.9b) we will have

(3.14)
$$
\begin{aligned}
X^1(t) &= X^1(0) + W^1(t) + B^1(t) + (1 - p_{11})Y^1(t) - U^{01}(t) + U^1(t) \\
X^2(t) &= X^2(0) + W^2(t) + B^2(t) + (1 - p_{22})Y^2(t) - p_{12}Y^1(t) \\
&\quad - U^{02}(t) - U_c^{12}(t).
\end{aligned}
$$

The $Y_c^{12}(\cdot)$ can be obtained from the limit. $Y^1(\cdot)$ via (3.9). The limits $(1 - p_{11})Y^1(\cdot) = \lim_\epsilon (Y^{10,\epsilon}(\cdot) + Y^{12,\epsilon}(\cdot))$ and $(1 - p_{22})Y^2(\cdot) = \lim_\epsilon Y^{20,\epsilon}(\cdot)$ are to be obtained from the limit of (3.13). Furthermore, (as in [1]) the $(1-p_{ii})Y^i(\cdot)$ obtained from the limits in (3.13) are the unique continuous functions which can increase only when $X^i(t)$ is zero and which guarantee that $X^i(t) \geq 0$.

The $U_c^{12}(\cdot)$ can be used to define $U^{12}(\cdot)$ via limits in (3.9). We will have $U^{12}(\cdot) = p_{12}U^1(\cdot)/(1-p_{11})$.

4. Description of the Limit Control Problem

In this section, we define the proper limit control problem for the system of Figure 2.1. First, it will be convenient to picture the effects of various control actions on the $X^\epsilon(\cdot)$ for small ϵ. We do this in some detail, since the limit problem is somewhat non-standard, partly owing to the possibility of 'multiple simultaneous impulses'. Also, the set of admissible impulses and associated costs are defined via the possible limits of the controlled $X^\epsilon(\cdot)$, associated with bounded costs.

Given the limit controlled reflected diffusion $X(\cdot)$, we will need to determine an optimal or δ-optimal policy for it. In order for the 'limit' problem to make sense, for any admissible policy π for the limit $X(\cdot)$, there must be a sequence π^ϵ of policies which can be applied to the $X^\epsilon(\cdot)$ (i.e., P_{ij}, P_i *on/off or rate controls) and such that, under* π^ϵ, $X^\epsilon(\cdot)$ *converges to* $X(\cdot)$ (with policy π), and the associated costs also converge. Because of this, the limit control problem must be defined in terms of limits of what is possible for the $X^\epsilon(\cdot)$. This yields a rather interesting limit control problem.

Controls for the Limit Problem. Refer to Figure 4.1, where some typical paths are constructed, under the heavy traffic conditions. Start at point (a) with all P_i, P_{ij} on except that P_{01} is off. The path moves to the left and as $\epsilon \to 0$, it converges to the horizontal line (a,b). The mean (interpolated) movement to the left in time Δ is $g_{a1}\Delta/\sqrt{\epsilon} + 0(\Delta)$. Hence in the limit, as $\epsilon \to 0$, there is an impulsive change.

Now, restart at (d) with only P_{12} off. The path drops, and as $\epsilon \to 0$ it tends to the vertical line (d,e). In time Δ, the mean drop is

$p_{12}g_{d1}\Delta/\sqrt{\epsilon} + 0(\Delta)$. The same path is followed if only P_{02} is off or if P_1 and P_{01} are both off, although the 'drop' speed will be different. Now, restart at (e) with only P_1 off. The path moves toward (f) (for small ϵ), and the limit slope can be calculated from

$$\text{(4.1)} \qquad \frac{\text{net mean flow into } P_2}{\text{net mean flow into } P_1} = \frac{g_{a2} - (1-p_{22})g_{d2}}{g_{a1}} = -\frac{p_{12}g_{d1}}{g_{a1}}$$

If the path reaches (f), then P_{01} must be turned off. If, at (g), we turn P_1 back on (but leave P_{01} off), then the path moves toward (h). The effects of both P_1 and P_{12} being off simultaneously are the same as for P_1 being off alone. Over small intervals of length Δ, the $\tilde{A}$, $\tilde{D}$ and Y terms in (3.10) contribute very little to the paths (compared to the effects of the control actions), since they converge weakly to continuous functions.

Now refer to (i), and let only P_{01} and P_{02} be off. Then the path moves to (j) with a limit slope calculated as in (4.1) and yielding the slope

$$\text{(4.2)} \qquad [(1 - p_{22})g_{d2} - p_{12}g_{d1}]/(1 - p_{11})g_{d1}$$

Similarly, if only P_{01} and P_{12} are off at (i), then the path moves toward (j) with a limit slope

$$\text{(4.3)} \qquad [(1 - p_{22})g_{d2} - g_{a2}]/(1 - p_{11})g_{d1}$$

All finite sequences of arbitrary lengths of the impulses described in connection with Figure 4.1 are possible. Suppose (e) $\to$ (f) $\to$ (g) $\to$ (h). Then as $\epsilon \to 0$, it would appear that the limit $X(\cdot)$ jumps from (e) to (h) directly. But this (e) $\to$ (h) impulse must be realized as a concatenation of the basic

impulses described above. In general the limit control is specified by a sequence of off/on actions for the P_i, P_{ij}, in a *specified order*, and with the impulsive distance travelled between successive ('simultaneous') control actions specified. The cost paid for the impulses is precisely the impulsive costs defined by (2.9). The described limitation on the ways in which the impulses for $X(\cdot)$ can be created is important, if the control problem for the limit $X(\cdot)$ is to be properly related to that for $X^\epsilon(\cdot)$. In Section 6, we show that the problem can be quite tractable from a numerical point of view.

The instantaneous changes in the $U^\alpha(\cdot)$ can be readily read off from the limit sequences of simultaneous impulses. For illustration, we do it for the (e,f,g,h) sequence of Figure 4.1. Let e_i, etc. denote the i^{th} coordinate of the point (e), and let δU^α denote the increment in U^α. On (e,f), $\delta U^{10} + \delta U^{12} = f_1 - e_1$, $\delta U_c^{12} = e_2 - f_2$. On (f,g), $\delta U^{01} = \delta U^{10} + \delta U^{12}$, and the value is unimportant, since their effects cancel in (2.8a). Also, $\delta U_c^{12} = f_2 - g_2$. On (g,h), $\delta U^{01} = g_1 - h_1$. All non-specified δU^α are zero. The δU^{1i} always occur as $(\delta U^{10} + \delta U^{12})$.

The Limit Dynamical System. The Wiener Process. The limit system will be (3.14). It will turn out that the limit $W^i(\cdot)$ can be decomposed as follows (using the limits of the three terms in (3.10) which are used to define the $W^{i,\epsilon}(\cdot)$).

$$W^1(\cdot) = \tilde{A}^1(\cdot) + W_d^1(\cdot),\ W_d^1(\cdot) = -\tilde{D}^{10}(\cdot) - \tilde{D}^{12}(\cdot)$$

$$W^2(\cdot) = \tilde{A}^2(\cdot) + W_d^2(\cdot),\ W_d^2(\cdot) = -\tilde{D}^{20}(\cdot) + \tilde{D}^{12}(\cdot).$$

Here, all the terms are continuous martingales, with $\tilde{A}^1(\cdot)$, $\tilde{A}^2(\cdot)$, $\tilde{D}^{20}(\cdot)$ and $(\tilde{D}^{10}(\cdot),\tilde{D}^{12}(\cdot))$ being mutually orthogonal. The quadratic variation of $\tilde{A}^i(\cdot)$

is $\int_0^t g_{ai}^3 \ \sigma_{ai}^2(X(s))ds$ and that of $W_d(\cdot) = (W_d^1(\cdot),\ W_d^2(\cdot))$ is $\Sigma(t) = \{\Sigma_{ij}(t)\}$, where

$$
(4.4) \qquad
\begin{aligned}
\Sigma_{11}(t) &= g_{d1}[p_{11}(1 - p_{11})t + g_{d1}^2(1 - p_{11})^2\int_0^t \sigma_{d1}^2(X(s))ds] \\
\Sigma_{12}(t) &= -g_{d1}^3 p_{12}(1 - p_{11}) \int_0^t \sigma_{d1}^2(X(s))ds - p_{12}p_{11}g_{d1}t \\
\Sigma_{22}(t) &= g_{d2}[p_{20}(1 - p_{20})t + p_{20}^2 g_{d2}^2 \int_0^t \sigma_{d2}^2(X(s))ds] \\
&\quad + g_{d1}[p_{12}(1 - p_{12})t + p_{12}^2 g_{d1}^2 \int_0^t \sigma_{d1}^2(X(s))ds].
\end{aligned}
$$

If the σ_{di}^2 and σ_{ai}^2 are constants, then the covariance is precisely that obtained by Reiman [1] (with a different notation used there).

It is evident from (4.4) and the cited orthogonality properties that there are mutually independent Wiener processes $w_a^i(\cdot)$, $w_d^i(\cdot)$, $w_d^{20}(\cdot)$, $\{w_d^{11}(\cdot),\ w_d^{12}(\cdot)\}$, where each scalar valued process is standard, and with respect to which $X(\cdot)$ is non-anticipative and $Ew_d^{11}(t)w_d^{12}(t) = -[p_{11}p_{12}/(1-p_{11})\ (1-p_{12})]^{1/2}t$ and

$$
(4.5) \qquad
\begin{aligned}
\tilde{A}^i(t) &= g_{ai}^{3/2} \int_0^t \sigma_{ai}\ (X(s))dw_a^i(s) \\
W_d^1(t) &= [g_{d1}p_{11}(1 - p_{11})]^{1/2} w_d^{11}(t) + (1 - p_{11})g_{d1}^{3/2} \int_0^t \sigma_{d1}(X(s))dw_d^1(s) \\
W_d^2(t) &= [g_{d2}p_{20}(1 - p_{20})]^{1/2} w_d^{20}(t) + [g_{d1}p_{12}(1 - p_{12})]^{1/2} w_d^{12}(t) + \\
&\quad + p_{20}g_{d2}^{3/2} \int_0^t \sigma_{d2}(X(s))dw_d^2(s) \\
&\quad - p_{12}g_{d1}^{3/2} \int_0^t \sigma_{d1}(X(s))dw_d^1(s).
\end{aligned}
$$

The terms involving $w_d^{ij}(\cdot)$ are due to the variations in the routing,

whereas the terms involving $w_d^2(\cdot)$ are due to variations in the service times.

The drift terms $B^i(\cdot)$ in (3.14) came from (3.6) and (3.7) and are

$$B^1(t) = \int_0^t [a^1(X(s)) - (1 - p_{11})d^1(X(s))]ds$$

$$B^2(t) = \int_0^t [a^2(X(s)) - (1 - p_{22})d^2(X(s)) + p_{12}d^1(X(s))]\,ds.$$

Then the limit problem is defined by (3.14).

Admissible Control Actions. The U_c^α and U^α in (3.14) are non-decreasing piecewise constant functions which have only a finite number of jumps on each finite interval, and they can be taken to be right continuous. They thus correspond to 'impulsive' controls. We first identify the allowed control impulses in the limit model (3.14) with those described above for the discrete model (2.8). The allowed impulsive effects of U^1 in (3.14) are those described for $U^{1,\epsilon}$ in (2.8), as $\epsilon \to 0$. Also the impulsive effects of U_c^{12} are the limits of those of $U_c^{12,\epsilon}$, and the effects of the U^{0i} are those of the $U^{0i,\epsilon}$ as $\epsilon \to 0$. This completely characterizes the possibilities for the impulse control of (3.14). Generally, several components of the controls might jump simultaneously, or a single jump in one component might be a consequence of a multiple simultaneous off/on sequence. We must allow these possibilites and distinguish an order for the 'simultaneity', as discussed above, not only because they are possible control actions, but because they are possible limits of control actions for the physical processes. Thus, we count the parts of the multiple simultaneous impulses as distinct impulses. We now develop the notation for keeping track of the necessary information. Recall the definitions of τ_n^ϵ and R_n^ϵ given below (2.9).

Let τ_n denote the sequence of event times. The τ_n are not necessarily

distinct, but $\tau_{n+1} \geq \tau_n$ and the subscript n denotes the correct ordering, 'simultaneous' or not. At each event time one or more of P_i or P_{ij} might shut off or on. What happens is indicated by the vector $R_n = (R_n^{01}, R_n^1, R_n^{02}, R_n^{12})$, where $R_n^{ij} = 1, -1$ or 0 (resp., R_n^1) according to whether or not P_{ij} (resp., P_1) is turned on, off or not changed at τ_n. *Associated with* (τ_n, R_n) *is* $\delta U_n = (\delta U_n^{01}, \delta U_n^{02}, \delta U_n^1, \delta U_{c,n}^{12})$, *the instantaneous (at* τ_n*) change in the controls* $U(\cdot)$. To illustrate the procedure refer to the path (e,f,g,h) in Figure 4.1. There are four event times, τ_1 associated with (e) τ_2 with (f), etc. Also $\tau_1 = \tau_2 = \tau_3 = \tau_4$. At τ_1, $R^1 = 1$. At τ_2, $R^{01} = 1$. At τ_3, $R^1 = -1$ and at τ_4, $R^{01} = -1$. All non listed R^α are zero. The associated impulses δU_n are given in the discussion below (4.3).

The $\{\delta U_n, \tau_n, R_n\}$ is said to be a control policy. The policy is said to be *admissible* if the function

$$\hat{\mathcal{R}}(t) = \{X_0, \delta U_n I_{\{\tau_n \leq t\}}, \tau_n I_{\{\tau_n \leq t\}}, R_n I_{\{\tau_n \leq t\}}, I_{\{\tau_n \leq t\}}, n < \infty, X(t), Y(t)\} \tag{4.7}$$

is non-anticipative with respect to the Wiener processes $w_\beta^\alpha(\cdot)$. An equivalent definition of admissibility is if the $\tilde{A}^i$, $\tilde{D}^{ij}(\cdot)$ are martingales with respect to the filtration generated by $\{\hat{\mathcal{R}}(t), \tilde{A}^i(\cdot), \tilde{D}^{ij}(\cdot)\}$, with the quadratic variation defined in and above (4.4).

Given $W(\cdot)$, $B(\cdot)$, $U(\cdot)$, there are unique processes $X(\cdot)$ and $Y(\cdot)$ such that $Y_i(\cdot)$ increases only when $X_i(t) = 0$, and where $X_i(t) \geq 0$ and (3.14) holds, as in [1] (see the end of Section 3). Of course, here $B(\cdot)$ and $w(\cdot)$ might depend on $X(\cdot)$, so it is not known a-priori that (4.6) has a unique solution. If the $\sigma_{ai}, \sigma_{di}, b^i$ do not depend on x, then the situation (without controls) is like that in [1] and we do have uniqueness of the solution to (3.14)

for each admissible contrc' policy. The $Y(\cdot)$ in (3.14) is obtained from (5.1) below, which is in turn obtained by taking limits in (3.13). In (5.1), $\{\mu_n^1\}$ is the subset of $\{\tau_n\}$ at which or both P_1 and P_{01} are on, with at least one turned off at τ_{n-1}, and $\{\mu_n^2\}$ is the subset of times at which all of P_1, P_{12} and P_{02} are on, with at least one being off at τ_{n-1}.

For an admissible policy, the cost function (the limit of (2.9)) is

$$
\begin{aligned}
V(\pi,x,P) = {} & E_x^\pi \int_0^\infty e^{-\beta t} k(X(t))dt + k_1 E_x^\pi \sum_n e^{-\beta v_n^i} \\
& + \sum_1^2 k_{0i} E_x^\pi \sum_n e^{-\beta v_n^{0i}} + k_{12} E_x^\pi \sum_n e^{-\beta v_n^{12}} \\
& + E_x^\pi \int_0^\infty e^{-\beta t} \left[\sum_{i=1}^2 q_{0i}\, dU^{0i}(t) + q_{12} d[U_c^{12}(t) - U^{12}(t)] \right]
\end{aligned}
\tag{4.8}
$$

In (4.8), the v_n^{ij}, v_n^i are defined as the moments of shutting off/on the indicated links or processors, as in Section 2.

5. Weak Convergence

We will use

A5.1. *The uncontrolled* $X(\cdot)$ *has a unique solution (in the weak sense) for each initial condition.*

Note that (A5.1) implies weak uniqueness of the solution $X(\cdot)$ for any admissible control policy.

Theorem 5.1. *Assume* (A2.1) *to* (A2.6) *and* (A5.1), *and let* $\sup_\epsilon V^\epsilon(\pi^\epsilon, X_0^\epsilon) < \infty$, *for* $\pi^\epsilon = \{R_n^\epsilon, \tau_n^\epsilon, \delta U_n^\epsilon, n < \infty\}$ *admissible. Then*

$$\{\tilde{A}^{1,\epsilon}(\cdot), \tilde{A}^{2,\epsilon}(\cdot), (\tilde{D}^{10,\epsilon}(\cdot), \tilde{D}^{12,\epsilon}(\cdot)), \tilde{D}^{20,\epsilon}(\cdot)\}$$

is tight in $D^5[0,\infty)$ *(Skorohod topology) and the limits of any weakly convergent subsequence of the four sets (we pair* $\tilde{D}^{10}$ *and* $\tilde{D}^{12}$*) are orthogonal continuous martingales. On each* $[0,t]$, *the mean number of control actions is finite, and the set of intervals on which some control is active converges to a finite set of points. The pieces*(+) *of* $X^\epsilon(\cdot)$ *on the intervals where no controls are active are tight, and the weak limits of these 'pieces' are continuous. The convergences* (3.9b) *all hold.*

Let ϵ *index a weakly convergent subsequence of* $\mathcal{R}^\epsilon = \{X_0^\epsilon, \tilde{A}^{i,\epsilon}(\cdot), \tilde{D}^{ij,\epsilon}(\cdot), B^\epsilon(\cdot), R_n^\epsilon, \tau_n^\epsilon, \delta U_n^\epsilon, i, ij, n\}$ *with limit denoted by* $\mathcal{R}$. *Define the process* $\mathcal{R}(\cdot)$ *from the limit processes by*

$$\mathcal{R}(t) = (X_0, \tilde{A}^i(t), \tilde{D}^{ij}(t), B(t), i, j, (R_n, \tau_n, \delta U_n) I_{\{\tau_n \le t\}}, n < \infty).$$

(+) More precisely, define the 'pieces' by shifting the start of the intervals to the origin, and continuing the 'piece' to the right of the interval by setting its value there to be equal to the value at the right end point of the interval.

Then $\tilde{A}^i(\cdot)$ *and* $\tilde{D}^{ij}(\cdot)$ *are martingales on the filtration engendered by the* $\mathcal{R}(t)$, *with the quadratic variations given in and above* (4.4). *The limit policy* $\pi = \{R_n, \tau_n, \delta U_n\}$ *is admissible for* $X(\cdot)$. *Except at points where there is control action,* (3.14) *holds, where* $Y^i(\cdot)$ *is defined by* (5.1). *(See* (3.13)*). We define (w.l.o.g.)* $X(t)$ *by* (3.14) *even at points of control action.*

$$\begin{aligned}&((1-p_{11})Y^1(\cdot),(1-p_{22})Y^2(\cdot)) = \\ (5.1)\qquad & \qquad F(W(\cdot),B(\cdot),0,X_0,X^i(\mu_n^i),\mu_n^i,\mu_n^i,\ i=1,2,\ n<\infty)\ .\end{aligned}$$

In (5.1) μ_n^i *is the limit of both* $\mu_n^{i,\epsilon}$ *and* $\hat{\mu}_n^{i,\epsilon}$ *and* $X^1(\mu_n^i)$ *is the limit of the values of* $X^{i,\epsilon}(\tilde{\mu}_n^{i,\epsilon})$ *(the* $\{\mu_n^i\}$ *are obtainable from the* $\{\tau_n,R_n\}$*). The* $Y^i(\cdot)$ *increase only when* $X^i(t) = 0$. *The limits of the uncontrolled sections of* $X^\epsilon(\cdot)$ *do not depend on the the subsequence, except for their initial conditions.*

Proof. (a) First, we show the convergences (3.9b). We do it for $\tilde{U}^{10,\epsilon}(\cdot) = U^{10,\epsilon}(\cdot) - p_{10}U^{1,\epsilon}(\cdot)/(1-p_{11})$ only, for the rest are treated in the same way. We have that

$$\tilde{U}^{10,\epsilon}(t) = \sqrt{\epsilon}\sum_1^{t/\epsilon}\frac{[I_n^{10}p_{12} - I_n^{12}p_{10}]}{(p_{10}+p_{12})}\psi_n^1\,(1-P_n^1)$$

is a martingale and its variance is bounded by $O(\epsilon)\ E\sum_1^{t/\epsilon}(1-P_n^1) = C^\epsilon(t)$. It is easily seen that

$$\overline{\lim_\epsilon}\ \sqrt{\epsilon}\ \sum_1^{t/\epsilon}(1-P_n^1) < \infty,$$

for otherwise the buffer of P_1 will fill up (one or more times), forcing the P_{01} to shut off (one or more times) such that $EU^{01,\epsilon}(t)$ will diverge and the costs will go to infinity as $\epsilon \to 0$. Thus, $C^\epsilon(t) \longrightarrow 0$ as $\epsilon \longrightarrow 0$, which yields the desired result.

By (3.9b), the $\rho^{i,\epsilon}(\cdot)$ and $\hat{\rho}^{i,\epsilon}(\cdot)$ of (3.10), (3.11) go to zero.

Below, the tightness of $\{W^{\epsilon}(\cdot)\}$ will be shown, together with the fact that its limits are continuous. This and the representation (3.13) implies that (for any weakly convergent subsequence) the $Y^{i,\epsilon}(\cdot)$ converge (in the Skorohod topology) to continuous processes $Y^i(\cdot)$. Thus, via (3.9), we have $Y_c^{12}(\cdot) = p_{12}Y^1(\cdot)$.

(b). We have $\epsilon S^{i,\epsilon}_{\alpha,t/\epsilon}$ and $\bar{S}^{i,\epsilon}_{\alpha}(t)$ converging weakly to the processes ($S^i_{\alpha}(\cdot)$ and $\bar{S}^i_{\alpha}(\cdot)$, resp.) with values $t/g_{\alpha i}$ and $g_{\alpha i}t$, resp. This is more or less obvious since (e.g.) $\epsilon \sum_1^{t/\epsilon} [\alpha^{i,\epsilon}_n - \bar{\alpha}^{i,\epsilon}_n]$ has orthogonal increments and its variance tends to zero as $\epsilon \longrightarrow 0$. The increments of each $\tilde{A}^{i,\epsilon}_0(\cdot)$ and $\tilde{D}^{ij,\epsilon}_0(\cdot)$ are also orthogonal. Due to the uniform integrability in (A2.3), those processes are tight and all weak limits are continuous martingales.

The four elements of ($\tilde{D}^{10}$ and $\tilde{D}^{12}$ are paired) $(\tilde{A}^{1,\epsilon}_0(\cdot), \tilde{A}^{2,\epsilon}_0(\cdot), (\tilde{D}^{10,\epsilon}_0(\cdot), \tilde{D}^{12,\epsilon}_0(\cdot)), \tilde{D}^{20,\epsilon}_0(\cdot))$ are mutually orthogonal, and so are the weak limits. To see the mutual orthogonality, one uses a calculation of which the following is typical. Take a 'typical' term from $\tilde{A}^{i,\epsilon}_0(\cdot)$ and $\tilde{D}^{ij,\epsilon}_0(\cdot)$ and use the definition of $E^{i,\epsilon}_{\alpha,n}$ above (A2.2) and the centering in (3.1) to get (drop the ϵ for simplicity)

$$
\begin{aligned}
&E[I^{ij}_k - p_{ij}\, \Delta^i_k/\bar{\Delta}^i_k][1 - \alpha^i_n/\bar{\alpha}^i_n] \\
&\quad = E[I^{ij}_k - p_{ij}\, \Delta^i_k/\bar{\Delta}^i_k] I_{\{S^{i,\epsilon}_{d,k-1} \le S^{i,\epsilon}_{a,n-1}\}} E^{i,\epsilon}_{a,n-1}(1 - \alpha^i_n/\bar{\alpha}^i_n) \\
&\quad + E[1 - \alpha^i_n/\bar{\alpha}^i_n] I_{\{S^{i,\epsilon}_{a,n-1} < S^{i,\epsilon}_{d,k-1}\}} E^{i,\epsilon}_{d,k-1}(I^{ij}_k - p_{ij}\, \Delta^i_k/\bar{\Delta}^i_k) \\
&\quad = 0 .
\end{aligned}
$$

Using the results in the first part of (b) above, and the definitions of $\tilde{A}^{i,\epsilon}(\cdot)$ and $\tilde{D}^{ij,\epsilon}(\cdot)$, all weak limits of $\tilde{A}^{1,\epsilon}(\cdot)$, $\tilde{A}^{2,\epsilon}(\cdot)$, $(\tilde{D}^{10,\epsilon}(\cdot), \tilde{D}^{12,\epsilon}(\cdot))$, $\tilde{D}^{20,\epsilon}(\cdot)$ are continuous martingales. All the assertions of the theorem

(except for the non-anticipativity assertion and the quadratic variation values) follow from the results in part (a) and (b) above.

(c) Owing to the mutual orthogonality of the four processes $\tilde{A}^{i,\epsilon}(\cdot)$, etc., and to (A2.3), we can calculate (for the limit process) the quadratic variation and prove the martingale property with respect to the σ-algebra engendered by $\mathcal{R}(\cdot)$ separately for each component. We do it only for $(\tilde{D}^{10,\epsilon}(\cdot),\tilde{D}^{12,\epsilon}(\cdot))$. Let ϵ index a weakly convergent subsequence of $\mathcal{R}^{\epsilon}$ and define $\mathcal{R}(\cdot)$ as in the theorem statement. Let $f(\cdot)$ be a smooth function with compact support and $h(\cdot)$ a bounded and continuous function, both real valued. Let $t,t+s$ and $t_k \leq t$ below be points such that the probability

$$P(\tau_n \text{ equals } t \text{ or } t+s \text{ or } t_k) = 0$$

for each n,k. Define $\delta\psi_n^{ij,\epsilon} = I_n^{ij,\epsilon} - p_{ij}\Delta_n^{i,\epsilon}/\bar{\Delta}_n^{i,\epsilon}$.

By the uniform integrability (A2.3), the representation of $\tilde{D}^{ij,\epsilon}(\cdot)$ as a sum, and a truncated Taylor series expansion, we can write (we can assume w.l.o.g. that $\tau_n^{\epsilon} \leq t$ for only finitely many n)

$$\begin{aligned}(5.2)\qquad & Eh\left(X^{\epsilon}(t_k),\tilde{A}^{i,\epsilon}(t_k),\tilde{D}^{ij,\epsilon}(t_k),B^{i,\epsilon}(t_k),(R_n^{\epsilon},\tau_n^{\epsilon},\delta\bar{U}_n^{\epsilon})I_{\{\tau_n^{\epsilon}\leq t_k\}},\ k,n\right)\cdot \\ & \Bigg[f(\tilde{D}^{10,\epsilon}(t+s),\tilde{D}^{12,\epsilon}(t+s)) - f(\tilde{D}^{10,\epsilon}(t),\tilde{D}^{12,\epsilon}(t)) \\ & - \sqrt{\epsilon}\sum_{\alpha=0,2}\ \sum_{\epsilon n=\bar{S}_d^{1,\epsilon}(t)}^{\bar{S}_d^{1,\epsilon}(t+s)} f_{x_\alpha}\left[\sqrt{\epsilon}\sum_1^{n-1}\delta\psi_k^{10},\sqrt{\epsilon}\sum_1^{n-1}\delta\psi_k^{12}\right]\cdot\delta\psi_n^{1\alpha} \\ & - \frac{1}{2}\epsilon\sum_{\alpha,\beta=0,2}\ \sum_{\epsilon n=\bar{S}_d^{1,\epsilon}(t)}^{\bar{S}_d^{1,\epsilon}(t+s)} f_{x_\alpha x_\beta}\left[\sqrt{\epsilon}\sum_1^{n-1}\delta\psi_k^{10},\sqrt{\epsilon}\sum_1^{n-1}\delta\psi_k^{12}\right]\cdot\delta\psi_n^{1\alpha}\,\delta\psi_n^{1\beta}\Bigg] \\ & \xrightarrow{\epsilon} 0 .\end{aligned}$$

Now, use the definition of $E^{1,\epsilon}_{d,n}$ given above (A2.2), the centering of $\delta\psi^{ij,\epsilon}_{k}$ and the assumption (A2.6) on the conditional variances to replace $\delta\psi^{1\alpha}_{n}$ in the first sum in (5.2) by zero and the $\delta\psi^{1\alpha}_{n}\delta^{1\beta}_{n}$ in the second by $E^{1,\epsilon}_{d,n-1}\delta\psi^{1\alpha}_{n}\delta\psi^{1\beta}_{n}$. This latter quantity is

$$E^{1,\epsilon}_{d,k-1}\left[I^{1\alpha}_{k} - p_{1\alpha}\frac{\Delta^1_k}{\bar{\Delta}^1_k}\right]\left[I^{1\beta}_{k} - p_{1\beta}\frac{\Delta^1_k}{\bar{\Delta}^1_k}\right]$$

$$= p_{1\alpha}\delta_{\alpha\beta} - p_{1\alpha}p_{1\beta} + p_{1\alpha}p_{1\beta}\,\mathrm{var}\,\Delta^1_k/(\bar{\Delta}^1_k)^2$$

(5.3)

$$= p_{1\alpha}\delta_{\alpha\beta} - p_{1\alpha}p_{1\beta} + p_{1\alpha}p_{1\beta}\,g^2_{d1}\sigma^2_{d1}(X^{\epsilon}_{S^{1,\epsilon}_{d,k-1}})$$

$$+ \text{(negligible terms)}\,.$$

The limit (as $\epsilon \longrightarrow 0$) of the double sum in (5.2) is

$$\frac{1}{2}\sum_{\alpha,\beta=0,2}\int_{\bar{S}^1_d(t)}^{\bar{S}^1_d(t+s)} f_{x_\alpha x_\beta}(\tilde{D}^{10}(\tau),\tilde{D}^{12}(\tau))\cdot\hat{\Sigma}_{\alpha\beta}(\tau)\,d\tau$$

where

$$\hat{\Sigma}_{00}(t) = p_{10} - p^2_{10} + p^2_{10}g^2_{d1}\sigma^2_{d1}(X(t/g_{d1}))$$

$$\hat{\Sigma}_{02}(t) = -\,p_{10}p_{12} + p_{10}p_{12}g^2_{d1}\sigma^2_{d1}(X(t/g_{d1}))$$

$$\hat{\Sigma}_{22}(t) = p_{20} - p^2_{20} + p^2_{20}g^2_{d1}\sigma^2_{d1}(X(t/g_{d1}))\,,$$

where we used (5.3) and the fact that $\epsilon S^{1,\epsilon}_{d,t/\epsilon} \longrightarrow t/g_{d1}$ to get the proper limit of the argument of $\sigma^2_{d1}(\cdot)$.

Now, recalling that $\bar{S}^i_d(t) = g_{di}t$, and taking limits in (5.2) yields

$$
\begin{aligned}
& Eh(X(t_k),\tilde{A}^i(t_k),\tilde{D}^{ij}(t_k),B^i(t_k),(R_n,\tau_n,\delta U_n)I_{\{\tau_n \le t_k\}}, n,k)) \cdot \\
(5.4)\qquad & \Bigg[f(\tilde{D}^{10}(t+s),\tilde{D}^{12}(t+s)) - f(\tilde{D}^{10}(\tau),\tilde{D}^{12}(t)) \\
& - \frac{1}{2}\sum_{\alpha,\beta=0,2} \int_{tg_{di}}^{(t+s)g_{di}} f_{x_\alpha x_\beta}(\tilde{D}^{10}(\tau),\tilde{D}^{12}(\tau))\ \hat{\Sigma}_{\alpha\beta}(\tau)d\tau \Bigg] = 0 .
\end{aligned}
$$

The arbitrariness of $h(\cdot)$, $f(\cdot)$, and t , $t+s$, $\{t_i\}$ (possibly excluding a countable set) imply that $(\tilde{D}^{10}(\cdot),\tilde{D}^{12}(\cdot))$ is a martingale with respect to the asserted filtration.

The quadratic variation can be obtained from (5.4) via a change of variables and is $\int_0^t \bar{\Sigma}(\tau)d\tau$, where $\bar{\Sigma}(\cdot) = \{\bar{\Sigma}_{\alpha\beta}(\cdot),\alpha,\beta = 0,2\}$ and

$$\bar{\Sigma}_{00}(t) = g_{d1}[p_{10} - p_{10}^2 + p_{10}^2\sigma_{d1}^2(X(t))]$$

$$\bar{\Sigma}_{02}(t) = g_{d1}[-p_{10}p_{12} + p_{10}p_{12}g_{d1}^2\sigma_{d1}^2(X(t))]$$

$$\bar{\Sigma}_{22}(t) = g_{d1}[p_{20} - p_{20}^2 + p_{20}^2\sigma_{d2}^2(X(t))] .$$

With analogous calculations for $\tilde{D}^{20,\epsilon}(\cdot)$ and for the $\tilde{A}^{i,\epsilon}(\cdot)$, we get quadratic variation for the $W^i_\alpha(\cdot),\tilde{A}^i(\cdot)$, $\tilde{D}^{i,j}(\cdot)$,as given in Section 4.

By the above argument the limit policy $\{\tau_n,R_n,\delta U_n\}$ is 'non-anticipative' with respect to the martingales, or their generating Wiener processes $w^\alpha_\beta(\cdot)$. Owing to the way they were obtained as limits of the $\{\tau^\epsilon_n,R^\epsilon_n,\delta U^\epsilon_n\}$, the limit policy $\{\tau_n,R_n,\delta U_n\}$ is admissible in the sense that it corresponds to admissible

sequences of impulses corresponding to the sequence of off/on controls as discussed in Section 4.

By the above argument, the limit policy $\{\tau_n, R_n, \delta U_n\}$ is 'non-anticipative' with respect to the martingales or their generating Wiener processes.

Q.E.D.

Extension. Consider the graph of $X^{\epsilon}(\cdot)$ ($X^{1,\epsilon}(\cdot)$ plotted vs. $X^{2,\epsilon}(\cdot)$) in the state space during a fixed control action. It can be shown that the graph converges uniformly (in probability) to the limit straight lines given by Figure 4.1, or the considerations leading to it in other cases. The convergence is in the sense that the maximum value of the distance between any point on (this part of) the graph of $X^{\epsilon}(\cdot)$ and the closest point on the limit straight line goes to zero in probability.

Theorem 5.2. *Assume* (A2.1) *to* (A2.6) *and* (A5.1), *and let* ϵ *index a weakly convergent subsequence with limit* $\mathfrak{X}(\cdot)$. *Then (with* π *defined as in Theorem* 5.1*) for any* P

(5.5) $$\varliminf_{\epsilon} V^{\epsilon}(\pi^{\epsilon},x,P) \geq V(\pi,x,P).$$

Define $N^{\alpha,\epsilon}(t)$ *to be the number of actions of the control* P_{α} *on the interval* $[0,t]$. *If*

(5.6) $$\{N^{\alpha,\epsilon}(n+1) - N^{\alpha,\epsilon}(n),\ \alpha,\ n < \infty\}$$

is uniformly integrable, then

(5.7) $$V^{\epsilon}(\pi^{\epsilon},x,P) \longrightarrow V(\pi,x,P).$$

Proof. The relation (5.5) is just a consequence of Fatous' Lemma and the weak convergence. Now, let the uniform integrability hold. Then, certainly the holding costs and the impulsive control costs in (2.9) converge to their limits, as given by the terms in (4.8). We need only work with the last integral in (2.9). The arguments for each component are essentially the same, and we work with the $U^{01,\epsilon}(\cdot)$ term only assuming that P_1 is on. If P_1 might also be off part of the time, the argument is a little more involved (involving the $X^{2,\epsilon}$ as well as the $X^{1,\epsilon}$), but is essentially the same.

When P_{01} is off, the increments in the $Y^{ij,\epsilon}(\cdot)$ are zero. [If $X^{1,\epsilon}(t) = 0$, we must have P_{01} on, by (A2.1)]. We can write

$$
\begin{aligned}
U^{01,\epsilon}(t) &= \sum_n [U^{01,\epsilon}(\bar{v}_n^{01,\epsilon} \cap t) - U^{01,\epsilon}(v_n^{01,\epsilon} \cap t)] \\
&= \sum_n [W^{1,\epsilon}(\bar{v}_n^{01,\epsilon} \cap t) - W^{1,\epsilon}(v_n^{01,\epsilon} \cap t)] \\
&\quad - \sum_n [X^{1,\epsilon}(\bar{v}_n^{01,\epsilon} \cap t) - X^{1,\epsilon}(v_n^{01,\epsilon} \cap t)] + \sum_n [B^{1,\epsilon}(\bar{v}_n^{01,\epsilon} \cap t) - B^{1,\epsilon}(v_n^{01,\epsilon} \cap t)]. \\
&\qquad + (\text{terms which } \to 0 \text{ as } \epsilon \to 0).
\end{aligned}
$$

For some $K_1 < \infty$, the last two sums on the right are bounded by $K_1 N^{01,\epsilon}(t)$, which is uniformly integrable by hypothesis. By the orthogonality properties of the summands in the expression for the $W^{1,\epsilon}(\cdot)$, the mean square value of the middle term is $O(t+1)$. This yields the uniform integrability of $\{U^{01,\epsilon}(t)\}$ for each t and of $\{U^{01,\epsilon}(n+1) - U^{01,\epsilon}(n),\ \epsilon > 0,\ n < \infty\}$. By the weak convergence and the uniform integrability of these and the other terms in the last integral of (2.9), the assertion (5.7) follows.

Q.E.D.

It is not a priori obvious that there is a control policy for which (5.6) is uniformly integrable, since we must shut off the inputs to P_i whenever its buffer is full. We will define a standard 'comparison' control policy called the Δ_0-*boundary policy*. It will be useful since its properties imply that we can always assume the uniform integrability of (5.6) for the optimal or δ-optimal policies for the $X^{\epsilon}(\cdot)$. Let $\Delta_0 \in (0,\min(B_1,B_2)/4)$ and refer to Figure 5.1. If $X^{2,\epsilon} = B_2$ then shut off all inputs to P_2 until $X^{2,\epsilon}$ reaches $B_2 - \Delta_0$. Then turn them back on. If at the end of that time $B_1 - \Delta_0 < X^{1,\epsilon} \leq B_1$, shut P_{01} off until $X^{1,\epsilon} = B_1 - \Delta_0$. If $X^{1,\epsilon} = B_1$, then shut P_{01} off until $X^{1,\epsilon}$ reaches $B_1 - \Delta_0$. Then turn P_{01} back on. We use the analogous definition for the Δ_0-boundary policy for $X(\cdot)$. Then, if ever $X^{\epsilon}(\cdot)$ or $X(\cdot)$ hits the outer boundary, we control it to a distance at least Δ_0 (in each coordinate) from the outer boundary.

Theorem 5.3. *Assume* (A2.2) *to* (A2.6). *Then for the* Δ_0*-boundary control and each* $k < \infty$

$$\sup_{\substack{\epsilon \text{ small} \\ \alpha,x,n}} E_x\left|N^{\alpha,\epsilon}(n+1) - N^{\alpha,\epsilon}(n)\right|^k < \infty, \text{ all } \alpha, \tag{5.8}$$

and similarly for the 'jump numbers' of the limit process $X(\cdot)$.

Remark on the proof. Refer to Figure 5.1. Let t_i^{ϵ} denote the i^{th} time of return of $X^{\epsilon}(\cdot)$ to the outer boundary after the i^{th} time that the control takes the process to the set $[0,(B_1 - \Delta_0)] \times [0,(B_2 - \Delta_0)]$. One can readily show that for any $\delta_0 \in (0,1)$, there is $T_0 > 0$ such that

(5.9) $$\sup_{\substack{\omega,i,\\ \text{small } \epsilon}} P\{t_{i+1}^{\epsilon} - t_i^{\epsilon} < T_0 \mid \text{data up to } t_i^{\epsilon}\} \leq 1 - \delta_0 .$$

This is just a consequence of the properties of $W^{\epsilon}(\cdot)$, $B^{\epsilon}(\cdot)$ and of the fact that $dU^{\alpha,\epsilon}(\cdot) = 0$ on the intervals of interest. With (5.9), it is not hard to show that all the moments of $N^{\alpha,\epsilon}(iT_0+T_0) - N^{\alpha,\epsilon}(iT_0)$ are bounded, uniformly in i and ϵ and in the initial condition. (Similarly, for the $X(\cdot)$ process.) This yields the desired result. See the proof of Theorem 5.3 in [7] of a related result for a problem with a more complicated statistical structure.

The optimality and 'almost' optimality theorem. At the present time almost nothing is known about optimal or δ-optimal ($\delta>0$) policies for the $X^{\epsilon}(\cdot)$. This is one of the basic reasons for considering suitably adapted policies which are 'good' for $X(\cdot)$. Unfortunately, we know little about the optimal or δ-optimal policies for $X(\cdot)$. Thus, we must postulate (in (A5.2)) the existence of a δ-optimal policy with certain smoothness properties. The assumption appears to be eminently reasonable, since there is usually enormous flexibility in the smoothing that can be put on δ-optimal controls. The numerical results obtained via the methods described in Section 6 satisfy (A5.2) for all the cases tried, in the sense that the 'control decision' surfaces (discretized for the numerical calculation) seem to have the required properties. In fact, the situation in Figure 5.1 is more or less typical, in the sense that some continuous deformation of these decision surfaces is usually the case.

For our current purposes, it is best to view the path $X(\cdot)$ as its graph in the state space. The uncontrolled sections are the graphs of the paths of the uncontrolled reflected diffusion, and the controlled sections are straight lines, each one (or perhaps part of one) correspond to a different value of the set of

indicators $P = (P^{01}, P^{02}, P^1, P^{12})$. In a sense, (A5.2) is a long-winded and formal way of saying that the lengths of the straight line segments are piecewise continuous in their starting point. It also deals with the possibility that the initial P might be inappropriate for the initial state x, and that we might have to change the control settings instantaneously at $t = 0$. We tried to give a general description of what reasonably seems to be expected. The situation might be simpler in special cases - but it seems likely that the useful δ-optimal (or even optimal) control policies would be described by (A5.2), due to the nature of the impulse sequences. Note that (the k_α are the cost coefficients in (2.6))

$$1 + \sup_{x,P} [V(x,P) + 1]/\min_{\alpha} k_\alpha \equiv \bar{K}$$

is an upper bound for the numer of 'simultaneous impulses' (the above number of sequential line segments) for the δ-optimal controls, with $\delta \leq 1$. We know that $\sup_{x,P} V(x,P) < \infty$, owing to the properties of the comparison Δ_0-boundary control of Figure 5.1.

We require some 'smoothness' in the δ-optimal 'feedback' controls, since we need to adapt them for use with the $X^\epsilon(\cdot)$ process and will require that the corresponding sequence $\{X^\epsilon(\cdot)\}$ (and the associated costs) converge appropriately to $X(\cdot)$ (and its associated cost).

The boundaries of the sets $G(1)$ and $G_i(P)$ below are smooth in that they are composed of a finite number of differentiable curves which are not tangent at the points of intersection. We use P to denote the control value just before a decision to change the control is made, and P_1 to denote the new control value just after the decision is made. Recall that $P = 1$ is used for $P = (1,1,1,1)$.

We could replace (A5.2) by the simpler assumption that for each $\delta > 0$ there is a δ-optimal admissible policy π_δ for $X(\cdot)$ and admissible policies π_δ^ϵ for $X^\epsilon(\cdot)$ such that $X^\epsilon(\cdot)$ (under π_δ^ϵ) $\Rightarrow X(\cdot)$ (under π_δ), and the associated costs converge. (A5.2) simply defines a reasonable π_δ for which this can be done. The interiors of all sets in (A5.2) are relative to $G = [0,B_1] \times [0,B_2]$.

A5.2. *For each* $\delta > 0$, *there is a 'feedback' policy* π_δ *for* $X(\cdot)$ *which is* δ-*optimal in the sense that it satisfies* (A2.1) *and*

$$(5.10) \qquad V(x,P) = \inf_{\pi \text{adm.}} V(\pi,x,P) \geq V(\pi_\delta,x,P) - \delta$$

for all x,P *and which has the following properties.*

(a) *Let* $P = 1$. *Then there is a decision set* $G(1)$, *whose boundary is divided into a finite number of segments. Each segment is associated with a switch to some* $P_1 \neq 1$ *when* $X(\cdot)$ *hits it from the outside. The segment associated with each* P_1 *is strictly interior to one of the sets* $G_i(P_1)$ *below.*

(b) *For each* $P \neq 1$, *there are a finite number (perhaps zero - see remark in* (c) *below) of sets* $G_i(P)$ *whose interiors are disjoint. If* $x \in G_i(P)$ *and* P *is used, then it is used until the boundary of* $G_i(P)$ *is reached. The distance (taken by the graph of* $X(\cdot)$, *which is a straight line) from* $x \in G_i(P)$ *to the boundary of* $G_i(P)$ *is a continuous function of* x. *The (straight line) graph is (uniformly) not tangent to the boundary at any point of contact. The boundary is divided into a finite number of segments, each associated with a new control setting, perhaps with* $P = 1$.

These segments are strictly interior to some set $G_j(P_1)$ *for the new value* P_1.

At the corners of the segments of $\partial G_i(P)$ *or* $\partial G(1)$, *any policy associated with the intersecting segments can be used. There is* $\Delta_1 > 0$ *such that after a finite number of switches, we have* $P = 1$ *and* $X(\cdot)$ *is a distance* $\geq \Delta_1$ *from* $G(1)$.

(c) *It is possible that there will be an immediate change* $P \to$ *some* $P_1 \neq P$ *at* $t = 0$. *If this occurs, we want the line segment of the graph of* $X(\cdot)$ *after the switch to correspond to* P_1 *for at least a minimum distance, independent of* x. *(This seems to be rather unrestrictive). We formalize this as follows.*

(c_1) *If we do not switch at* $t = 0$, *then assume that* $x \in$ *some* $G_i(P)$ *above.*

(c_2) *If we do switch (to some* $P_1 \neq P$*) at* $t = 0$. *Then assume that* $x \in$ *some* $G_i(P_1)$ *above and* $\inf_{x \in G_i(P_1)} d[x, \partial G_i(P_1)] > 0$.

Remark. The assumption concerning 'points in common' to several $\partial G_i(P)$ does not seem to be restrictive. Generally, in dynamic programming, when the state is on the boundary of sets corresponding to different policies, any one of the policies is optimal. Condition (A5.2) is intended to be illustrative of the possibilities that we can allow.

Adapting π_δ **to** $X^\epsilon(\cdot)$. By adapting the policy π_δ for use with $X^\epsilon(\cdot)$ we simply take as the moments of decision the moments when $X^\epsilon(\cdot)$ hits the decision boundary segments.

We now prove the 'almost δ-optimality' of π_δ-applied to $X^\epsilon(\cdot)$. Theorem 5.4 says essentially that a 'nice' control which is almost optimal for $X(\cdot)$ will also be almost optimal for $X^\epsilon(\cdot)$. This justifies the use of the limit approximations for purposes of getting good or nearly optimal controls.

Theorem 5.4. *Assume* (A2.2) *to* (A2.6), (A5.1) *and* (A5.2) . *Let* π_δ^ϵ *denote the policy of* (A5.2) *adapted to* $X^\epsilon(\cdot)$. *Then*

(5.11) $$V^\epsilon(\pi_\delta^\epsilon,x,P) \longrightarrow V(\pi_\delta,x,P)$$

uniformly in x . *For admissible* π^ϵ *and small* ϵ ,

(5.12) $$\sup_{\{\pi^\epsilon\}} \sup_x \,[V^\epsilon(\pi_\delta^\epsilon,x,P) - V^\epsilon(\pi^\epsilon,x,P)] \le 2\delta .$$

Proof. The proof is a consequence of the weak convergence in Theorems 5.1 and 5.3, the piecewise continuity properties of (A5.2) and an estimate of the type obtained in Theorem 5.2, and we only outline some of the argument.

(a) The facts that the segments of G(1) are piecewise differentiable with non-tangent corners and that the uncontrolled $X(\cdot)$ is non-degenerate imply that the hitting times (and locations) of $X^\epsilon(\cdot)$ on G(1) converge to those for the limit $X(\cdot)$, for any initial condition outside G(1).

(b) Similarly for the hitting times and locations of the boundaries of the $G_i(P)$, when $P \neq 1$.

(c) The uncontrolled segments of $X^\epsilon(\cdot)$ converge to those of $X(\cdot)$. The graphs of the controlled segments of $X^\epsilon(\cdot)$ converge uniformly to their limit straight line segments, as discussed in the remark after Theorem 5.1.

(d) If a limit point of $X^\epsilon(\cdot)$ or a limit point of an end point of a segment of the graph during a control interval - is on a corner of the boundary

of $G(1)$ or of some $G_i(P)$, then the limit control actions just after contact with the boundary there is, of course, specified by the limit of the control actions of $X^{\epsilon}(\cdot)$ just after contact with the boundary. But, by (A5.2) which of the actions associated with that boundary point are used for $X(\cdot)$is irrelevent. Whatever it is, it will be used for a positive minimum distance (on the graph).

(e) Let $N^{\epsilon}(t)$ denote the number of distinct control actions on $[0,t]$. Then a proof such as would be used to prove Theorem 5.3 together with the weak convergence and the fact that $\Delta_1 > 0$ can be used to show that $\{N^{\epsilon}(n+1) - N^{\epsilon}(n),\ n < \infty,\ \epsilon > 0\}$ is uniformly integrable. (This is then used as in Theorem 5.2.).

(f) Let ϵ index a weakly convergent subsequence. The limit process is the $X(\cdot)$ associated with π_{δ}. By (A5.1) and (A5.2), the particular sequence used is irrelevent.

(g) (5.11) follows from the above facts and theorems 5.1 and 5.2.

(h) (5.12) follows from the theorem 5.2 and the fact that π_{δ} is δ-optimal for $X(\cdot)$. The limits of the controls $\{\pi^{\epsilon}\}$ might depend on the subsequence. But (5.12) holds uniformly in the subsequence.

Extensions. The arrival and service time sequences can each be correlated, (e.g., service in 'random batches', etc.), provided that they satisfy suitable mixing conditions. If they are correlated and state dependent, then the 'first order perturbed test function method' of [5, Chapter 5] (see also [6]) can be adapted. It is possible to control the service or arrival rates (marginal a^i,d^i) also. Impulsive controls (hence piecewise constant rates) are easy to accomodate here.

Otherwise, one can introduce relaxed controls as in [6], writing (e.g.) the drift term as $\int_0^t b^i(X^\epsilon(s),\alpha)m_t(d\alpha)$ where $m_t(\cdot)$ is the measure associated with the relaxed control. We do need to maintain heavy traffic, of course. The variances can also be allowed to be control dependent. There is no problem allowing this 'impulsively', but for continuously controlled variances, there is still some uncertainty concerning the appropriate description of the limit problem.

For more general feedforward - branching networks, controlling the p_{ij} might also be of interest. One could use $p_{ij} = \bar{p}_{ij} + \sqrt{\epsilon}\ \delta p_{ij} + o(\sqrt{\epsilon})$. Then, when the 'principal terms' are cancelled in (3.5), we are left with an additional 0(1) term-depending on $\{\delta p_{ij}\}$, and this corresponds to an additional drift associated with the 'marginal' control of the routing. Various types of controlled priority service are possible - and might be the subject of a future paper. For example, the customers might fall into various priority classes which relate, for example, to service time distributions. We might control the priority service subject to holding costs depending on the priority.

The average cost per unit time problem is trickier, but one can adapt the scheme for the ergodic problem in [6]. Here $\{X^\epsilon(\cdot)$, vector of elapsed times since the last service completions or arrivals$\}$ would replace the vector $\{X^\epsilon(\cdot), \xi^\epsilon(\cdot)\}$ of [6]. Then, under appropriate ergodicity conditions concerning the δ-optimal processor, we can extend Theorem 5.4.

6. A Numerical Method for Approximating the Optimal Value Function and Control

The control problem defined by the cost (4.8), system (3.14) and the control actions described by the possibilities associated with the off/on impulses associated with the discussion about Figure 4.1 can be approximated by the numerical methods studied in [9]. The method in [9] involved a Markov chain (indexed by a 'finite difference' approximation parameter) approximation to the optimal continuous time problem. One then showed that the sequence of value functions for the chains converged to the optimal value function for the continuous parameter problem, and that suitable continuous parameter interpolations of the chain converge weakly to the optimal controlled continuous parameter process. The methods of [9] can be readily adapted to our problem, and only an outline will be given. The weak convergence methods used in [9] will have to be replaced by the methods here - owing to the reflection term, but the general idea is the same.

Let h be a finite-difference approximation parameter, and B_i be integral multiples of h. Let G_h denote the h-grid on $G = [0,B_1] \times [0,B_2]$. Define a_{ij} by $\Sigma_{ij}(t) = \int_0^t a_{ij}(X(s))ds$, and generally omit the x-argument in the $a_{ij}(\cdot)$ and $b^i(\cdot)$ below. For the Markov chain approximation, the status of the controls at any time is defined by the vector $P = (P^{01},P^{02},P^1,P^{12})$, where $P^\alpha = 1$ (resp., 0) denotes that the control is on (the link is operating normally) (resp., closed). Recall that, when $P = (1,1,1,1)$, we write $P = 1$.

Let $\{X_n^h\}$ denote the approximating Markov chain, and let x denote the canonical current state, y the canonical successor state and P_1 the canonical control which will be used at state x to bring the chain to the next state.

Define $X^h(\cdot)$, the interpolated process to be the right continuous piecewise constant process with interpolation intervals $\Delta t^h(x,P_1)$. Both these intervals and the transition probabilities $p^h(x,y/P_1)$ depend on the new chosen control as well as on the current state. If $P_1 \neq 1$, we use $\Delta t^h(x,P_1) = 0$; i.e., the interpolation interval has zero length. In this case, several steps of $\{X_n^h\}$ all occur simultaneously in the interpolation $X^h(\cdot)$. Define $Q_h(x) = 2[a_{11} + a_{22} - |a_{12}|] + h(|b^1| + |b^2|)$, and let $a_{ii} - |a_{12}| \geq 0$, $i = 1,2$. For $P_1 = 1$, we use $\Delta t^h(x,P_1) = h^2/Q_h(x)$.

We now define the transition probabilites $p^h(x,y|P_1)$ for the chain when $P_1 = 1$, for $x,y \in G_h$. Let e_i denote the unit vector in the i^{th} coordinate direction. We use

$$
(6.1) \qquad \begin{aligned}
p^h(x,x\pm e_ih|P_1 = 1) &= [a_{ii} - |a_{12}| + h(b^i)^\pm]/Q_h(x), \\
p^h(x,x+e_1h-e_2h|P_1 = 1) &= p^h(x,x-e_1h+e_2h|P_1 = 1) \\
&= |a_{12}|/Q_h(x).
\end{aligned}
$$

If some x^i (the i^{th} component of x) equals zero - then the transition probability (6.1) is modified as follows, as a concatenation of two transitions, the first being (6.1). For the second (the 'reflection') step, we distinguish two cases.

Case 1: The 'y' argument in the p^h in (6.1) is not in G_h, but $x^1 \neq 0$ or $y \neq x - e_1h + e_2h$. Then simply project (reflect) the process back to the nearest point in G_h.

Case 2: Let $x^1 = 0$ and $y = x - e_1h + e_2h$. Then the second transition is back to $y = x$, with a probability $p_{12}/(1 - p_{11})$ and back to $y = x + e_2h$ with probability $[1 - p_{12}/(1 - p_{11})]$. This step is to account for the $p_{12}Y^1$ term in (3.14).

If $P_1 = 1$ always, then $X^h(\cdot) \Rightarrow X(\cdot)$, uncontrolled and unreflected [9].

Let P denote the control used to get the current state x. The actual state for the problem is the pair (x,P), since the cost associated with the next transition depends on whether or not some element of the current control vector is changed. Let $K^h(x,P,P_1)$ denote the costs associated with the transition, when current state is x, and control P changes to P_1. For $P_1 = 1$ $K^h(x,P,1) = \Delta t^h(x,1)k(x)$, the holding cost only.

We now define some of the transition probabilities and costs when $P_1 \neq 1$. There are 15 possibilities, and only some typical ones will be described. These are constructed so that the limit (as $h \to 0$) of $X^h(\cdot)$ will be the reflected controlled $X(\cdot)$, and so that the associated costs for $X^h(\cdot)$ will also converge to that for $X(\cdot)$. Write $P = (P^{01},P^{02},P^{12},P^1)$, $P_1 = (P_1^{01},P_1^{02},P_1^{12},P_1^1)$.

Let $P_1^{01} = 0$, with other $P_1^\alpha = 1$. Then use $p^h(x,x-e_1h|P_1) = 1$ (by (A2.1), $x^1 > 0$ here) and $K^h(x,P,P_1) = q_{01}h + k_{01}I_{\{p^{01}=1,p_1^{01}=0\}}$. Now, let $P_1^{02} = 0$ with other $P_1^\alpha = 1$. Then $p^h(x,x-e_2h|P_1) = 1$ and $K^h(x,P,P_1) = q_{02}h + k_{02}I_{\{p^{02}=1,\ p_1^{02}=0\}}$. For $P_1^{12} = 0$ and other $P_1^\alpha = 1$, we have $p^h(x,x-e_2h|P_1) = 1$ and $K^h(x,P,P_1) = q_{12}h + k_{12}I_{\{p^{12}=1,p_1^{12}=0\}}$.

Now, let $P_1^1 = 0$ with other $P_1^\alpha = 1$. Let $p_{12}g_{d1} \leq g_{a1}$ (the reverse case is treated anologously) and refer to Figure 6.1. The line from x to (a) is the mean direction of the appropriate impulse, and its slope (see Section 4) is $[g_{a2} - (1 - p_{22})g_{d2}]/g_{a1} = -p_{12}g_{d1}/g_{a1}$. In order to 'simulate' this mean line, we use

$$p^h(x,x + e_1h - e_2h|P_1) = p_{12}g_{d1}/g_{a1} = 1 - p^h(x,x + e_1h|P_1).$$

The instantaneous cost is $K^h(x,P,P_1) = k_1I_{\{p^1=1,p_1^1=0\}}$.

Now, let $P_1^{12} = P_1^{02} = 0$ with all other $P_1^\alpha = 1$. Then $p^h(x,x - e_2h|P_1) = 1$. The 'impulsive' part of $K^h(x,P,P_1)$ is obvious, namely $k_{12}I_{\{p^{12}=1,p_1^{12}=0\}} +$

$k_{02}I_{\{p^{02}=1,p_1^{02}=0\}}$. But the 'opportunity' cost - that due to Z^{12} and U^{02} is less obvious. This is obtained from the relative rates at which $X^2(\cdot)$ decreases due to the effects of P_{12} and P_{02} (resp.) being off. This is (resp.) $p_{12}g_{d1}$ and g_{a2}. Thus we use the 'opportunity' cost

$$h[q_{12}p_{12}g_{d1} + q_{02}g_{02}]/(p_{12}g_{d1} + g_{a2}).$$

The $p^h(x,y|P_1)$ and $K^h(x,P,P_1)$ are calculated in a similar way for all the other possibilities.

The dynamic programming equation for our 'approximation' problem is

$$V^h(x,P) = \min_{P_1} [(\exp - \beta\Delta t^h(x,P_1))\sum_y p^h(x,y|P_1)V^h(y,P_1) + K^h(x,P,P_1)]. \tag{6.2}$$

The weak convergence methods of this paper can be used to show that $V^h(x,P) \to V(x,P) = \inf_{\pi_{adm}} V(x,\pi)$. It can be shown that, for each x there is an (ω,t)-dependent control such that the approximation methods (for the control) in [9, Chapter 9] can be used. For reasonable grid sizes, say 50 × 50, the numerical problem is quite tractable.

For the numerical problem, we do not need to duplicate the dynamics of the original system $X^\epsilon(\cdot)$, but we can use any controlled process which has the same controlled limit equation. See the book [9] for a fuller development of this computational point of view for a large class of more classical problems.

References

[1] M.I. Reiman, "Open queueing networks in heavy traffic", Math. of Operations Research, **9**, 1984, p. 441-458.

[2] J. Michael Harrison, "Brownian models of queueing networks with heterogeneous customer populations", Stanford Graduate School of Business Report, Sept., 1986. To appear in Proc. Inst. of Math. and Applications, Proc. on Stochastic Differential Equations, Springer, 1987.

[3] J. Michael Harrison, "The diffusion approximation for tandem queues in heavy traffic", Adv. Applied Prob., **10**, 1978, p. 886-905.

[4] A.J. Lemoine, "Networks of queues - a survey of weak convergence results", Management Science, **24**, 1978, p. 1175-1193.

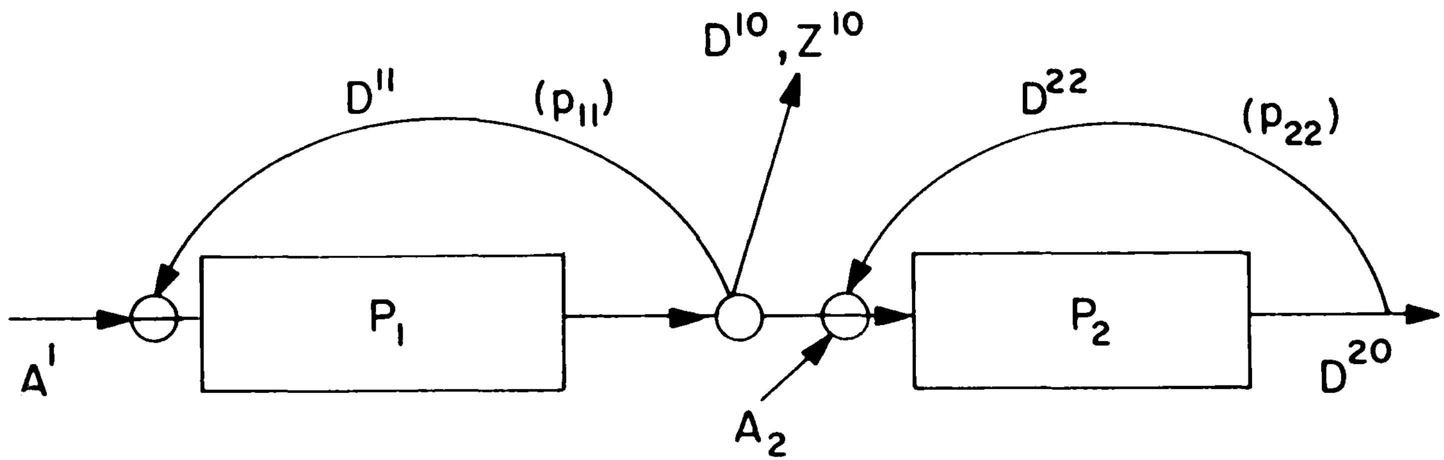
[5] H.J. Kushner, *Approximation and Weak Convergence Methods for Random Processes; with Application to Stochastic Systems Theory*, M.I.T. Press, Cambridge, Mass., USA, 1984.

[6] H.J. Kushner, and W. Runggaldier, "Filtering and control for wide bandwidth noise driven systems", IEEE Trans. on Automatic Control, AC-32, 1987, p. 123-133.

[7] H.J. Kushner, and K.M. Ramachandran, "Nearly optimal singular controls for wideband noise driven systems", LCDS Rept. 86-43, Brown Univ., submitted to SIAM J. on Control and Optimization.

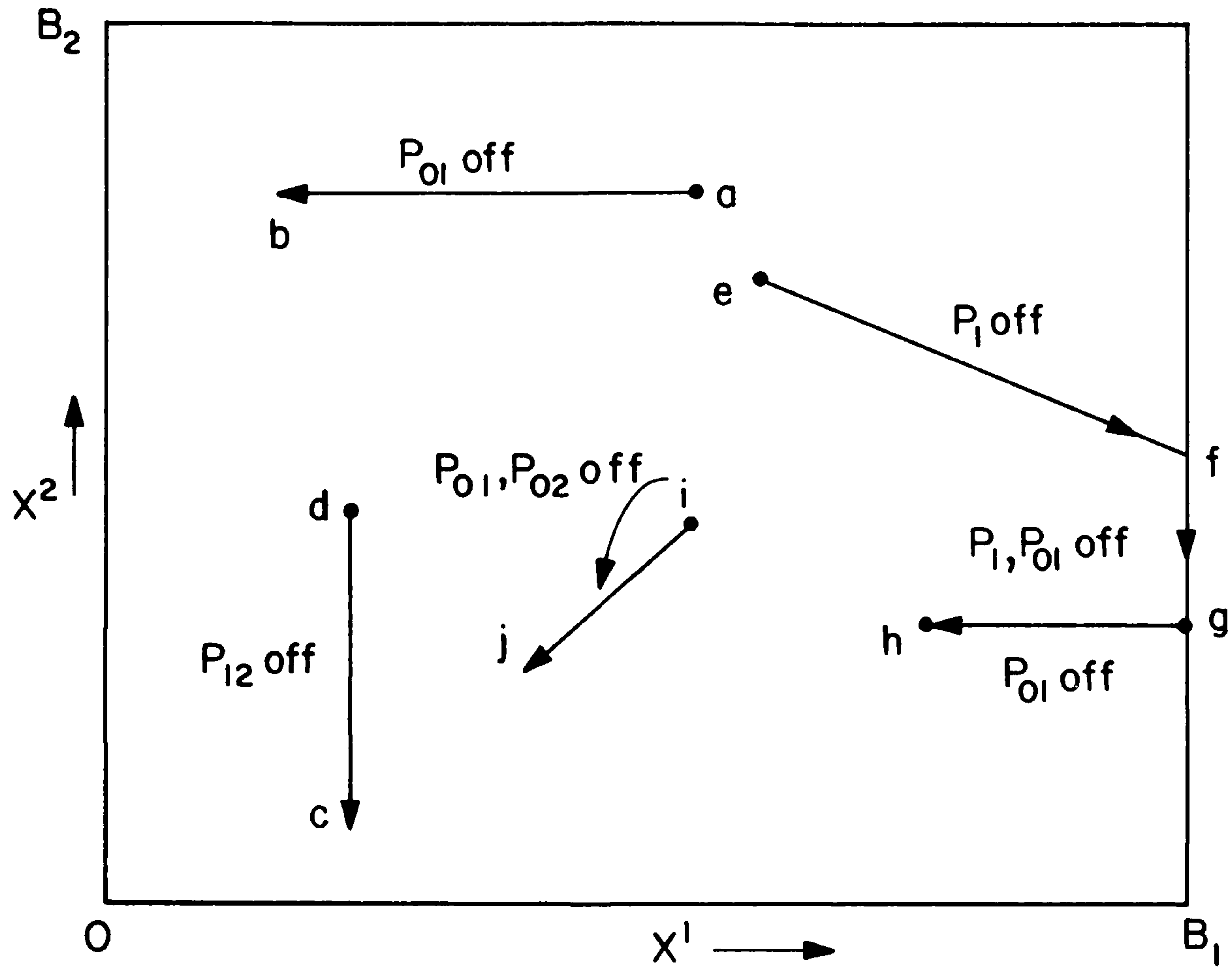
[8] H. Kushner, and W. Runggaldier, "Nearly optimal state feedback controls for stochastic systems with wideband noise disturbances", SIAM J. on Control and Optimization, **25**, 1987, p. 289-315.

[9] H.J. Kushner, *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, 1977, Academic Press, New York.



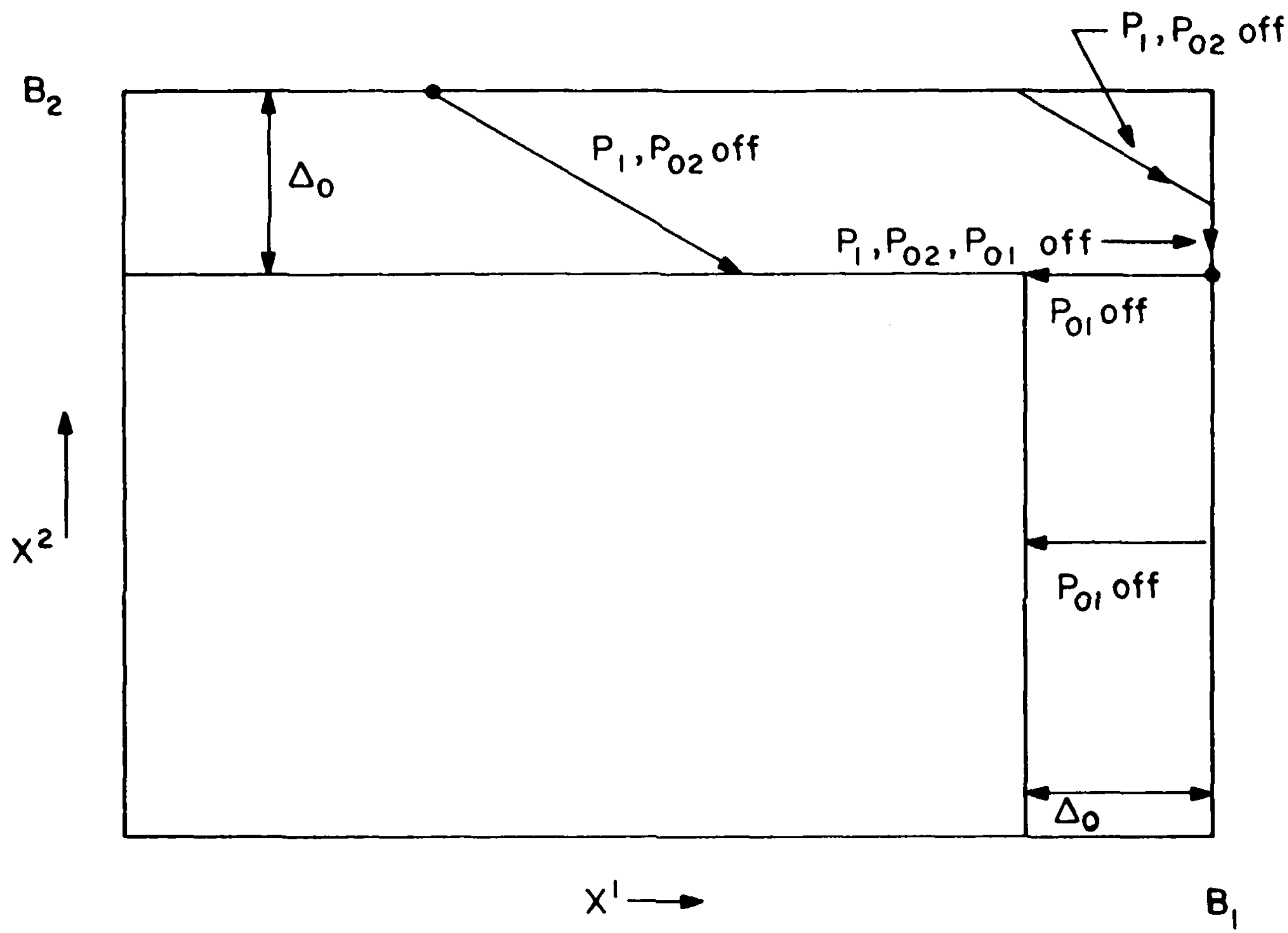
The System Configuration

Figure 2.1



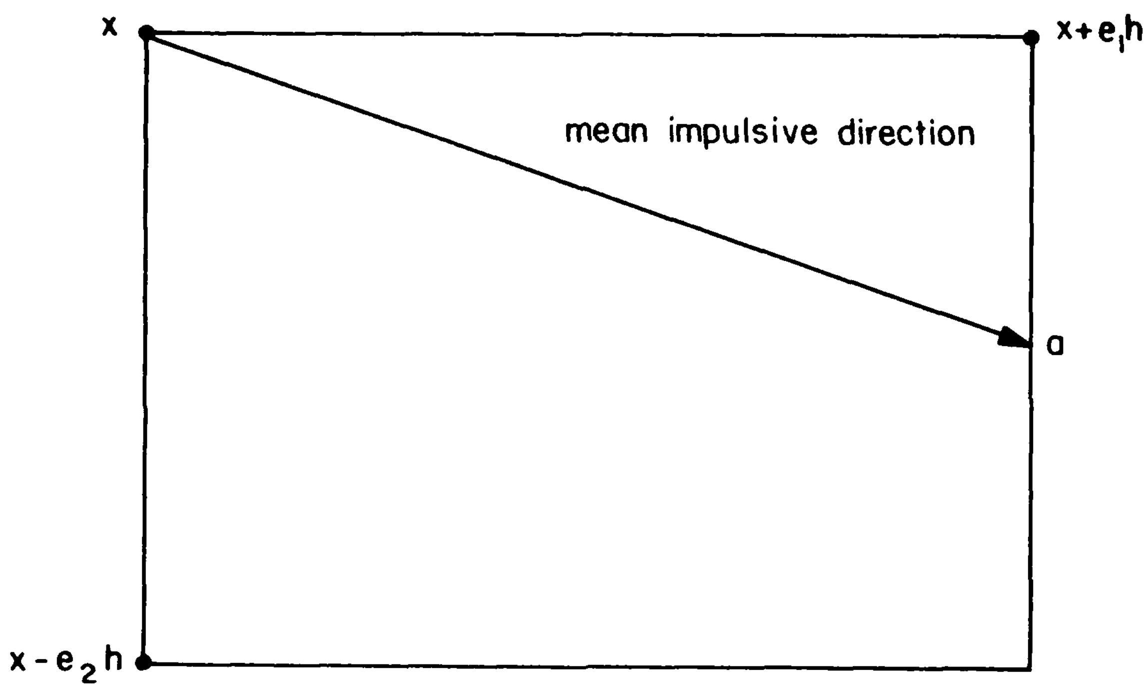
Impulsive Changes in $X^{\epsilon}(\cdot)$ or $X(\cdot)$ Due to the Control Actions

Figure 4.1



The Comparison Δ_0 - Boundary Control

Figure 5.1



$$P^h(x, x+e_1h-e_2h/\rho) = 1-P^h(x, x+e_1h/\rho) = p_{12}g_{d1}/g_{a1}$$

$$p_{12}g_{d1} \le g_{a1}$$

Transition Probabilities for $P' = 0$, other $P^{\alpha} = 1$

Figure 6.1

END
12-87
DTIC